AAMRL-TR-90-019

FIELD MEASUREMENT OF HEAD RELATED TRANSFER FUNCTIONS

FREDERIC WIGHTMAN, Ph.D.
DORIS J. KISTLER, Ph.D.

HEARING DEVELOPMENT RESEARCH LABORATORY
WAISMAN CENTER
UNIVERSITY OF WISCONSIN-MADISON

APRIL 1990

Final Report for Period May 1988 to August 1989

Approved for public release; distribution is unlimited

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
HUMAN SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-6573

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from Armstrong Aerospace Medical Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Center
Cameron Station
Alexandria, Virginia 22314

TECHNICAL REVIEW AND APPROVAL

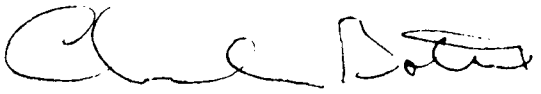
AAMRL-TR-90-019

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

The voluntary informed consent of the subjects used in this research was obtained as required by Air Force Regulation 169-3.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER



CHARLES BATES, JR.
Director, Human Engineering Division
Armstrong Aerospace Medical Research Laboratory

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704 0188

1a REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S) AAMRL-TR-90-019
6a NAME OF PERFORMING ORGANIZATION Board of Regents University of Wisconsin-Madison (6b OFFICE SYMBOL (If applicable))	7a NAME OF MONITORING ORGANIZATION Armstrong Aerospace Medical Research Laboratory (AAMRL/HEA)
6c ADDRESS (City, State, and ZIP Code) Madison WI 53706	7b ADDRESS (City, State, and ZIP Code) Wright-Patterson AFB OH 45433-6573
8a NAME OF FUNDING/SPONSORING ORGANIZATION (8b OFFICE SYMBOL (If applicable))	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 62202F; PROJECT NO: 7184; TASK NO: 26; WORK UNIT ACCESSION NO: 05

11 TITLE (Include Security Classification)

Field Measurement of Head-Related Transfer Functions

12 PERSONAL AUTHOR(S)
Wightman, Frederic, PhD and Kistler, Doris J., PhD

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM May 88 TO Aug 89	1989 August 15	173

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB GROUP	KEMAR Manikin; Auditory Signals; Anechoic Chamber; Digital Filters; Psychoacoustics; Psychophysical Assessment
23	06	-	
23	01	-	

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

This effort sought to refine and simplify techniques for generating acoustical signals that could be used in three-dimensional (3-D) auditory displays. Such signals are presented to a listener over headphones and create the illusion of a virtual sound source at a predetermined position in 3-D space. The signals are generated digitally, using algorithms based on the acoustical effects of human outer ear structures on sound waves reaching the ears. To date, the main area of difficulty inhibiting development of practical 3-D displays is in obtaining estimates of these outer ear effects. The focus of this effort was in this area.

The work was divided into three areas: 1) acoustical measurements of free-field-to-eardrum transfer functions (also called head-related transfer functions, or HRTFs); 2) analysis of HRTFs; and 3) psychophysical assessment of human performance in sound localization tasks involving stimuli presented both in real and in simulated (virtual) auditory space. The focus in all three areas was on evaluation of means for making HRTF measurements faster and easier, thus simplifying synthesis of auditory stimuli for 3-D displays. (continued on back)

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL Kenneth Gilleland, 1st Lt, USAF	22b TELEPHONE (Include Area Code) (513) 255-7902
	22c OFFICE SYMBOL AAMRL/HEA

DD Form 1473, JUN 86 Previous editions are obsolete SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

20. ABSTRACT - continued

In the measurement area, HRTFs were obtained from 20 human subjects at 144 positions in an anechoic sound field. A periodic pseudorandom noise averaging technique (Wightman and Kistler, 1989a) was used to make the measurements. Comparable HRTF measurements were also obtained from a KEMAR mannequin (using the same pseudorandom noise procedure) and from one of the original 20 subjects using a brief click as a measuring stimulus. The aim of obtaining HRTFs from KEMAR was to assess the need to base 3-D stimulus synthesis on individualized (listener specific) HRTF measurements. If acceptable measurements could be obtained from KEMAR, the time-consuming and somewhat risky measurement procedures involving real subjects could be eliminated. The motivation for the click measurements was to evaluate the feasibility of making HRTF measurements in an ordinary room, with appropriate gating to remove echoes.

Analysis of the HRTFs revealed large inter-subject differences, substantial differences between the KEMAR HRTFs and those from any of the human subjects, and a minimum of 20 dB loss in signal-to-noise (S/N) ratio accompanying the use of the click as a measuring stimulus. The magnitude component of the HRTFs from nearly all subjects included a deep notch, usually in the 8-12 kHz region, that was dependent on probe microphone position and independent of source direction. Further acoustical and optical measurements confirmed that this notch was a result of standing waves in the ear canal. A principal components analysis of the HRTFs was conducted with the aim of assessing the feasibility of constructing "model" HRTFs that would have the important features of real HRTFs. Unfortunately, available principal components algorithms do not accept complex data, so only the magnitude components of the HRTFs were analyzed. The analysis revealed that 90% of the variance in the HRTFs could be accounted for by 5 principal components. The first of these confirmed the overall similarity of the HRTFs across subjects in the low frequencies, and the next two revealed large differences across both subjects and positions in the important 5-15 kHz region.

Extensive psychophysical tests, using techniques developed and tested previously (Wightman and Kistler 1989b), were conducted on 15 adult listeners. In these tests, stimuli were presented from 36 positions either in free-field (anechoic chamber) or in simulated free-field (over headphones). Listeners gave numerical judgements of apparent azimuth and elevation of the sources. The results suggested: 1) when simulated free-field stimuli are synthesized from HRTF measurements obtained from the listeners' own ears, the apparent positions of the stimuli are the same as in free-field; 2) the elevation components of the apparent position judgements of simulated free-field stimuli are very sensitive to distortions (in the 5-10 kHz region) of the HRTFs used to synthesize the stimuli (such as occur from use of HRTFs from other listeners or from KEMAR). The tentative conclusions of the psychophysical tests were: 1) at the present time, the most veridical simulations of three-dimensional auditory space require synthesis to be based on a listener's own HRTFs; 2) because of the sensitivity of the apparent elevation of simulated sources, great care must be taken to preserve HRTF information in the 5-10 kHz region; 3) only if techniques can be developed which offer much higher S/N ratio than a single click will it be possible to obtain the necessary high-frequency detail in the HRTF measurements while making the measurements in an ordinary room.

PREFACE

The research described in this report was performed at the Hearing Development Research Laboratory, Waisman Center, University of Wisconsin-Madison. This research was supported by the Harry G. Armstrong Aerospace Medical Research Laboratory (AAMRL) Director's Funds. The project was performed under subcontract number 1107-02-07UWA with MacAulay-Brown, Inc., in support of prime contract number F33615-87-C-0534.

The principal investigators for this effort were Dr. Frederic Wightman and Dr. Doris J. Kistler. The Air Force project technical monitor was 1st Lt German Valencia of the Human Engineering Division, Visual Display Systems Branch (AAMRL/HEA).

Thanks are due to to Mr. Richard L. McKinley and Mr Mark A. Ericson of the Biodynamics and Bioengineering Division (AAMRL/BB). The authors also wish to express their appreciation to Ms. Gloria L. Calhoun (AAMRL/HEA) and Mr. Jeffrey R. Agnew, MacAulay-Brown, Inc. for reviewing the final report.

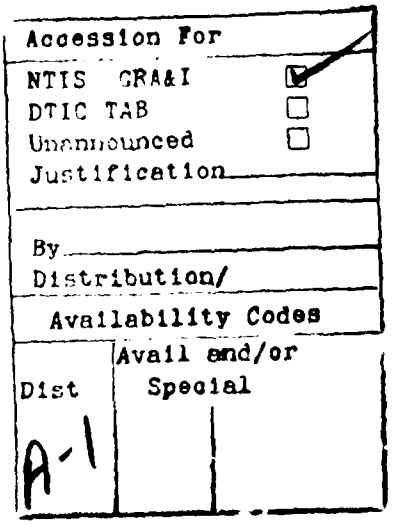


Table of Contents

LIST OF FIGURES

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

LIST OF TABLES

1.0 INTRODUCTION

A three-dimensional (3-D) auditory display has been identified as one of the virtual technologies associated with the Air Force "Super Cockpit" project. In addition to a panoramic visual display, the "Super Cockpit" will provide the pilot, via headphones, information from aircraft avionics, weapons, and navigation systems in a manner which optimizes the use of his spatial and psychomotor capabilities. The auditory display subsystem will allow the pilot to hear threats, targets and other operators as if they originated from specific locations in 3-D space. For example, verbal instructions from an electronic co-pilot will appear to originate from behind the pilot's head. These signals will be directionally accurate and stabilized in space regardless of the pilot's head position.

There have been very few extensively documented (i.e., with psychophysical data) demonstrations that 3-D auditory space can be successfully simulated with headphone-presented signals (see, for example, Wightman and Kistler, 1989b). However, it is generally agreed that veridical spatial simulation requires preprocessing of the signal, prior to headphone delivery, so as to mimic the acoustic effects of the head, shoulders, and outer ears. Such preprocessing is typically implemented in the form of a digital filter (one for each ear), the transfer function of which consists, in part, of an estimate of the acoustic free-field-to-ear-canal transfer function, or "head-related transfer function" (HRTF) as it is often called. Obtaining estimates of these transfer functions, in order to implement the digital filters required for spatial simulation, presents several significant problems. Finding solutions to these problems is the aim of this work.

The first problem arises because measuring the HRTFs is technically demanding and it is subject to numerous errors (Mehrgardt and Mellert, 1977; Wightman and Kistler, 1989a.). The degree to which these various sources of error contaminate the HRTF measurements in a perceptually significant way is not clear. For example, the measuring microphones are very small; thus, inherently noisy. Hence, positioning the microphone in a stable way in the ear canal is difficult, and since it must be close to the eardrum, there is some risk to the Subject. To reduce extraneous noise and echoes, the measurements should be made in a soundproof or ideally anechoic room. Another major problem arises from the fact that the extent of inter-individual differences in HRTFs across numerous subjects is not well-known. The possibility of potentially large differences (Wightman and Kistler, 1989a) suggests that the digital filters may have to be individual-specific for the simulations to be veridical. If true, this would complicate spatial simulation procedures enormously, since the HRTFs of each potential listener would have to be separately measured. Unless techniques could be developed to make such measurements in the field, to use a standard set of HRTF measurements for all listeners, or to model the HRTFs mathematically, simulation of auditory space via headphones could remain a laboratory curiosity.

Solution to the problems outlined above must come from research on both the engineering and psychophysical aspects of the issues. While an engineering approach can, for example, reveal the optimal technique for modeling HRTFs, only a psychophysical experiment can reveal the perceptual significance of

differences between the model HRTFs and the real HRTFs. Therefore, a combined approach was used for the work conducted under this effort. In parallel with developing new measurement techniques, psychophysical experiments were used to evaluate the perceptual consequences of various strategies for simplifying the measurements of HRTFs and thereby achieving a more immediate practical application.

The specific focus of the work was on the need for individualized HRTF measurements. The approach involved the following steps:

1. Measurement of HRTFs (both left and right ear) for sound sources at a large number (144) of positions, from a large number (20) of subjects, and from a standard mannequin (KEMAR), using well understood and proven techniques (Wightman and Kistler, 1989a.)

2. Analysis of the measured HRTFs to assess inter-individual variability in HRTF amplitude and phase characteristics in various frequency regions, and evaluation of analytic techniques (e.g., principal components) for reducing the HRTFs to weighted sums of underlying basis functions.

3. Psychophysical assessment, on a smaller number of 10, of the perceptual adequacy of auditory spatial simulations based on non-individualized HRTFs, HRTFs based on mannequin measurements, or, if the analysis is successful, "canonical" HRTFs synthesized on the basis of the multivariate analysis suggested above.

4. Assessment of the possible perceptual consequences of using HRTFs measured "in the field", in an ordinary room, with or without appropriate gating to remove echoes.

2.0 MEASUREMENT OF HRTFs

Our procedure for producing signals for a three-dimensional (3-D) auditory display involved digital synthesis of stimuli which are then presented over headphones. The basic assumption that guides this approach is that, if the acoustical waveforms at a listener's eardrums are the same under headphones as in free-field, the listener's perceptual experience will also be the same. Thus, we ignore the relevance of head movements, visual cues and other localization cues. However, the promising psychophysical results obtained to date (see Psychophysical Experiments section of this report, and Wightman and Kistler, 1989b) suggest that, for a limited range of listening conditions, the assumption is valid. The stimulus synthesis technique, the central feature of which is the measurement of free-field-to-eardrum acoustical transfer functions, is described in detail elsewhere (Wightman and Kistler, 1989a). That description is reprinted here for completeness and readability.

> Our approach is based on well-understood linear filtering principles. Let $x_1(t)$ represent an electrical signal which drives a loudspeaker in free-field, and let $y_1(t)$ represent the resultant electrical signal from a probe microphone positioned at a listener's eardrum. Similarly, let $x_2(t)$ represent an electrical signal which drives a headphone, with $y_2(t)$ the resultant microphone response. Given $x_1(t)$, our goal is to produce $x_2(t)$ such that $y_2(t)$ equals $y_1(t)$. We do this by designing a linear filter which transforms $x_1(t)$ into the desired $x_2(t)$.
>
> The design of the appropriate filter is best described in the frequency domain. Thus, $X_1(jw)$, or simply X_1, is the Fourier transform of $x_1(t)$, Y_1 is the transform of $y_1(t)$ and so forth. The probe microphone's response to $x_1(t)$ can be written:
>
> $$Y_1 = X_1 \, L \, F \, M \qquad (1)$$
>
> where L is the loudspeaker transfer function, F the free-field to eardrum transfer function (sometimes called the head-related transfer function, or HRTF), and M the microphone transfer function. The probe microphone's response to $x_2(t)$ can be written:
>
> $$Y_2 = X_2 \, H \, M \qquad (2)$$
>
> where H represents the headphone to eardrum transfer function. Setting $Y_1 = Y_2$ and solving for X_2 yields:
>
> $$X_2 = X_1 \, (L \, F \,) / H \qquad (3)$$
>
> This equation shows that the desired filter transfer function T is given by:

$$T = (L F)/H \quad (4)$$

Thus, if the signal, $x_1(t)$, is passed through this filter, and the resultant, $x_2(t)$, is transduced by the headphone, the signal recorded by the probe microphone at the eardrum will be $y_1(t)$, the same signal produced by the loudspeaker in free-field. This is represented in the frequency domain by substituting the right side of Equation (3) for X_2 in Equation (2).

The filter described in (4) applies only to a single free-field loudspeaker position and one ear. To synthesize each stimulus, then, we must design a pair of filters (one for each ear) for each desired free-field source position.

The first phase of our synthesis procedure involves measurement of the free-field-to-eardrum transfer function (HRTF) for each ear of a subject, for a large number of sound source positions. In practice, what we actually measure is a quantity like Y_1 in Equation (1) above, which includes not only the free-field-to-eardrum characteristics (F), but also the characteristics of the test signal (X_1), loudspeaker (L), and microphone (M). A headphone-to-eardrum transfer function (like Y_2 in Equation (2) above) is also measured for each ear of the same subject. In the second phase of the synthesis, each desired experimental stimulus is digitally filtered. The transfer functions of the filters (one for the left ear stimulus, and one for the right) are defined in Equation (4) above. Ideally, when the filtered stimuli are presented to the subject over the headphones, the waveforms reaching the eardrums should be identical to those produced by a free-field stimulus. The error in the procedure is quantified by recording the stimuli at the eardrums in the free-field and headphone conditions and computing the difference.

2.1 Transfer Function Measurement

Both free-field and headphone transfer function measurements were made using a technique loosely based on the procedure described by Mehrgardt and Mellert (1977). A wide-band, noise-like signal was presented (either by loudspeaker or headphone) repetitively, and the response at the listener's eardrum was obtained by averaging the output of a probe microphone. The Fast Fourier Transform (FFT) of this response was divided by the Fourier transform of the signal to produce an estimate of the transfer function in question. The signal was 20.48 msec in duration, and was computed via an inverse Discrete Fourier Transform (DFT) so that both the amplitude and phase components of its spectrum could be tailored to maximize the signal-to-noise ratio in the response

recordings. Specifically, the amplitude spectrum of the signal was flat from 200 Hz to 4000 Hz, where it increased abruptly by 20 dB. Thereafter, it was flat to 14 kHz. The signal contained no energy below 200 Hz or above 14 kHz. The phase spectrum was computed to minimize the peak factor of the signal (Schroeder, 1970). The signal was output continuously (hence with a repetition frequency of about 50 Hz), via a 16-bit digital-to-analog (D/A) converter (controlled by an IBM-PC) at a rate of 50 kHz. No anti-aliasing filters were used. For the free-field measurements, the signal was transduced by a miniature loudspeaker (Realistic Minimus-7). For the headphone measurements, the signal was transduced by a pair of Sennheiser HD-340 headphones, driven in phase. Signals were presented at approximately 70 dB SPL, a level chosen to reduce the contaminating effects of the acoustic reflex.

The acoustical response at the eardrum was measured with a miniature electret microphone (Etymotic) coupled to a silicone rubber probe tube with an outer diameter of less than 1 mm (see Figure 1). This probe microphone system, with its matching preamplifier and compensation network, had a sensitivity of about 50 mV/Pascal, and a frequency response which was relatively flat (+/- 2.5 dB) from 200 Hz to 14 kHz. Two matched microphones were used, one for each ear, and the responses from both were measured simultaneously. The amplified microphone outputs were digitized (simultaneously) using 16-bit analog-to-digital (A/D) converters (controlled by the IBM-PC) at a 50 kHz sampling rate. The responses to 1000 periods of the signal were averaged with floating-point precision, a spectral resolution of 48.8 Hz, and a worst-case signal-to-noise ratio of well over 20 dB in the range 200 Hz - 14 kHz.

The acoustical measurements were made with the tips of the probe-tubes positioned roughly in the middle of the subject's ear canal, about 1-2 mm from the eardrum. This position was chosen in order to be certain the measurements would capture all direction-dependent effects (which may not be the case for measurements at the ear-canal entrance) and to avoid standing-wave nulls at high frequencies. At 14 kHz, the highest frequency of interest in our work, the first standing wave null would occur at about 6 mm from the eardrum (assuming the ear canal is a uniform tube closed at one end). To avoid occluding the ear canals, the probe tubes were held in place with custom (i.e., different for each subject) lucite earmold shells, trimmed so that they did not extend into the concha when inserted, and bored out to a thickness of less than 0.5 mm. With the earmold shell in place, the probe tube was inserted into a thin, semi-rigid guide tube which was cemented to the wall of the earmold shell (see Figure 2). The length of each guide tube was calibrated, at the time the earmold assembly was made, so that with the probe inserted as far as its collar-stop would allow, the probe tip was about 1 mm from the eardrum. This calibration was accomplished by inserting a

human hair into the guide tube until the subject indicated that the hair had touched the eardrum. The hair was then marked and withdrawn so that the appropriate length for the guide tube could then be determined. The body of the microphone was left hanging at the side of the subject's ear.

For free-field measurements, the periodic wide-band signal was transduced by one of eight loudspeakers, each positioned 1.38 m from the subject in an anechoic chamber. The loudspeakers were mounted on a semicircular arc (2.76 m diameter), the ends of which were attached directly above and directly below the subject (see Figure 3). The loudspeakers were aimed at the position of the subject's head in order to minimize the influence of loudspeaker directionality (which we found to be virtually non-existent within 10 degrees of the speaker axis.) The entire arc assembly could be rotated (by hand crank) around the vertical axis, and positioned with a precision of about 0.5 degrees. The subject was seated on an adjustable stool (with back) so that his/her head was at the center of the arc. The speakers were mounted at -36, -18, 0, +18, +36, +54, +72, and +90 degrees elevation relative to the horizontal plane passing through the subject's ears. The measurements were made at all elevations except +72 and +90 degrees, and at all azimuths around the circle in 15 degree steps. Thus, transfer functions were measured from both ears at 144 source positions. Figure 4 shows a block diagram of the hardware used to make the HRTF measurements; the same set-up (without microphones) was used to produce the stimuli in the psychophysical experiments.

A typical measurement session lasted about an hour. After the microphones were fitted in the subject's ear canals, the subject was seated in the anechoic chamber, and instructed on how to set the azimuth of the loudspeaker speaker arc using the hand-crank to turn the arc. Then, with the subject alone in the chamber, the arc was moved to the first azimuth setting (usually directly behind the subject). Depending on the condition under study, the subject either looked directly forward and held his/her head still, or bit down on a bitebar, which could be attached rigidly to the subject's seat. After the subject signalled the experimenter that all was ready, measurements were made in rapid succession at all six elevations, in both ears simultaneously. About 2 minutes were required to make the six pairs of measurements at each azimuth. The subject then moved the arc to the next location and the sequence was repeated. Finally, after measurements had been made at all 24 azimuths, the subject put on the headphones, taking care not to disturb the position of the microphones, and a pair of transfer function measurements were taken with the headphones being used to transduce the wide-band test signal.

2.2 Digital Filter Construction

Each raw data record consisted of the time-domain representation of a signal recorded from a probe microphone in a subject's ear canal. This signal included not only the direction-specific characteristics of the subject's outer ear (and head, shoulders, etc.), but also the characteristics of the original test signal, the loudspeaker (or headphones), and the measuring microphone. To obtain an uncontaminated free-field-to-eardrum transfer function characteristic (HRTF) or an uncontaminated headphone-to-eardrum transfer function, the effects of the signal, loudspeaker (or headphone), and microphone must be removed. This could be done by transforming the raw data record into the frequency domain (via a FFT) and dividing by the frequency domain representation of the characteristics of the signal, the microphone, and the loudspeaker or headphone. In our case, to produce the digital filters required for stimulus synthesis, we divided the frequency domain representations of the signals recorded in free-field by the frequency domain representations of the same signals recorded under headphones. Since the stimulus and microphone characteristics appear in both the numerator and denominator terms, they cancel. The loudspeaker characteristics were not removed from the digital filters used to synthesize stimuli. All digital signal processing, including test stimulus generation, FFT computations, digital filter design and implementation, and waveform analysis was accomplished on a DEC VAX-11/750 computer using the ILS (Signal Technology Inc.) software package.

A complete set (144 positions) of HRTF measurements was obtained from 20 human subjects using the procedures described above. A representative subset of the measurements (all six source elevations, 12 azimuths) is shown for two subjects in Figures 5-28. In these figures, the magnitude response is shown on decibel coordinates, and the phase response (which has been "unwrapped" to avoid the usual ambiguities at + and - p boundaries) on radian coordinates. A full set of HRTF measurements was also obtained from KEMAR, using procedures that differed somewhat from those used with human subjects. First, the KEMAR we used had only one "ear" (pinna, canal model and microphone). Therefore, an assumption of symmetry around the vertical median plane was used to estimate HRTFs from the other ear. Second, KEMAR's own microphone (B&K 4134) was used to measure the HRTF. A representative sample of the HRTFs obtained from KEMAR is shown in Figures 29-40.

A second complete set of HRTF measurements was obtained from one of the 20 original subjects using a brief acoustical impulse (a 20 ms unipolar square pulse transduced by the loudspeakers) presented at a rate of about 50 per second as the measuring stimulus, in place of the usual periodic pseudorandom noise. All other aspects of the measuring procedure were the same. A representative sample of the measurements made with the click stimulus is shown in Figures 41-52.

3.0 ANALYSIS OF HRTFs

The HRTF data described in the previous section were analyzed in several different ways, in order to assess: 1) the variability, from subject to subject, in the amplitude and phase components of the HRTFs; 2) the sensitivity of the HRTF measurements to the position of the probe microphone in the subject's ear canal; 3) the differences between HRTFs obtained using the pseudorandom noise and HRTFs obtained using the click stimulus; and 4) the differences between HRTFs measured from human subjects and HRTFs measured from the KEMAR mannequin. In addition, a subset of the data was analyzed with two procedures, critical band smoothing and principal components, in an effort to develop rigorous procedures for representing HRTFs more simply.

3.1 Intersubject Variability in the HRTF

It has been known for over 20 years (e.g., Shaw, 1965) that large intersubject differences exist in the magnitude components of the HRTF. Our previous work (Wightman and Kistler, 1989a) quantified these differences and showed further that the differences are greatest in the 5-10 kHz region, and are not dependent on source position. A comparable analysis of the HRTF data gathered in this project produced similar results. Figures 53-56 show, for four source positions, the mean and 95% confidence intervals of the smoothed magnitude of the HRTFs from our 20 subjects. The magnitudes were smoothed using a critical bandwidth of 0.50. As reported before (Wightman and Kistler, 1989a), the intersubject variability in the magnitude response is greatest in the 5-12 kHz region, with 95% confidence intervals of 20 dB or more not uncommon.

Our approach to quantification of the intersubject differences in the phase components of the HRTF was based on the assumption that interaural phase differences would be most meaningful from a psychophysical point of view. Therefore, we ignored the monaural phase component of the HRTFs and examined interaural phase difference. More specifically, we examined intersubject differences in phase-derived interaural time difference, under the assumption that time difference is the more meaningful quantity from a perceptual standpoint. The interaural time difference vs frequency functions from these subjects are virtually identical to those we have published previously (Wightman and Kistler, 1989a), and the intersubject differences are nearly completely determined by head size.

3.2 Sensitivity of HRTF Measurements to Probe Microphone Position

Positioning a microphone probe tube in a subject's ear canal for HRTF measurements is a significant problem. The probe must be far enough down the ear canal, past the entrance, to capture all directional effects, close enough to the eardrum to avoid contamination of the HRTF by standing-wave nulls, but not so close to the eardrum as to risk injury. In addition, the probe must not move during the measurement or when headphones are worn. Our procedure attempts to solve the stability problem with the use of a custom earmold shell that holds the probe in place (Wightman and Kistler, 1989a). The probe is positioned by

inserting it a calibrated distance down a guide tube that is attached to the earmold shell.

The distance of the probe tip from the eardrum is critical, since the standing wave null in the measured HRTF response appears at a frequency inversely proportional to that distance (e.g., the null is at 8.5 kHz if the probe is 10 mm from the drum, and at 17 kHz if the probe is 5 mm from the drum). To avoid the standing wave null, and the resulting distortion and loss of HRTF information, we attempted to place the probe 2 mm or less from the eardrum (Wightman and Kistler, 1989a). The method we used relied on indirect estimates of probe-to-eardrum distance, relying on subjective report of when a human hair made contact with the eardrum. During the course of this project, we developed a more direct method for estimating the distance between the probe and the eardrum. The method is based on techniques described by Chan and Geisler (1989). First, the magnitude components of the HRTF measurements from a large number of source positions (144) are averaged, thus smoothing out the large direction-dependent spectral features. This average estimates the diffuse-field response of the ear (Shaw, 1980). Figures 57-59 show the diffuse-field response from three of our subjects. Note that in some cases, the diffuse-field response contains an obvious notch in the 6-12 kHz region. This notch is not present in Shaw's (1980) estimates of the diffuse-field response of the human ear, and is almost certainly a reflection of the standing-wave null. A second set of HRTF measurements is obtained with the microphone probe withdrawn a few millimeters. The magnitudes of these measurements are also averaged. The first average is divided by the second average. The common features of the diffuse-field response in the two averages cancel, leaving primarily the effects of the standing wave nulls in the two averages. One appears as a notch (from the first measurement, with the probe inserted at its maximum depth), the other (from the second measurement) appears as a peak. The frequencies of the notch and peak provide estimates of the probe-eardrum distances in the two cases. Figures 60-62 show the results of dividing the two diffuse-field estimates for 3 of our subjects. The probe-eardrum distance estimates derived from the two sets of HRTF measurements are shown in Table 1. Note that, for most of the subjects, the probe-eardrum distance is considerable greater than the expected 1-2 mm. This is most likely a result of the subjects not being able to determine when the human hair used as a depth probe actually touched the eardrum (Wightman and Kistler, 1989a). Table 1 also includes estimates of the probe-eardrum distance derived by combining an optical measurement of ear canal length with an estimate of entrance to probe tip distance taken from the earmold assembly. Note that, in most cases, the two probe-eardrum distance estimates are in good agreement.

We conclude from the experiments described above that HRTF measurements are very sensitive to probe position in the ear canal. The main reason appears to be a result of standing wave patterns in the earcanal. Unfortunately, even our best efforts to place the probe tip close to the eardrum (to avoid the standing wave problem) did not completely avoid the problems in all subjects. Since the standing wave nulls appear in the 5-12 kHz region, where important localization cues are believed to be encoded, it is possible that poor probe

placement for HRTF measurements could compromise the adequacy of 3-D simulations based on those measurements. The extent of the compromise will be evaluated in a subsequent section of this report.

3.3 Differences Between HRTFs Measured with Pseudorandom Noise and HRTFs Measured Using Clicks

The disadvantage of the pseudorandom noise signal used to measure HRTFs is that since it is periodic, room reflections cannot easily be removed, and thus, the measurements must be made in an anechoic room. With a transient such as a unipolar click, gating out reflections is a rather simple matter. It is theoretically possible, then, that HRTFs could be measured in an ordinary room if a click signal could be used. However, a click has considerably less energy than the pseudorandom noise stimulus, and thus, for a constant peak level, the S/N ratio in the measurements would be considerably poorer with a click signal. The peak level of the click cannot be raised to compensate for the loss in S/N ratio, since, at high levels the acoustic reflex contaminates the HRTF measurements (Wightman and Kistler, 1989a).

Figure 63 compares HRTFs obtained with the usual periodic pseudorandom noise (PRN) signal with HRTFs measured using a click, adjusted to maximum feasible level. Note that, while the click HRTFs have the same general shape as the PRN HRTFs, the magnitude is at least 20 dB less. At high frequencies, where the HRTF is normally reduced in magnitude, the click HRTF is rather different from the PRN HRTF, probably as a result of poor S/N ratio. The perceptual significance of these differences will be evaluated in a subsequent section of this report.

3.4 Differences Between HRTFs from KEMAR and HRTFs from Humans

Given the large intersubject differences in HRTF described above, we were led to expect large differences between HRTFs obtained from KEMAR and those obtained from human subjects. Figures 53-56 summarize those differences for four source positions. Note that while the KEMAR HRTFs generally fall within the range of HRTFs obtained from humans, there are consistent differences, the most striking of which is that the KEMAR HRTFs are higher in the high frequency regions. The effect of these differences on the adequacy of 3-D auditory simulations produced from the KEMAR measurements can only be determined from psychophysical experiments. The results of an experiment in which human listeners localized simulated sources produced from KEMAR's HRTFs are described in Section 4.4.

3.5 Simplification of HRTFs

The acoustical measurement procedures we used produce a very detailed representation of the HRTF. Impulse responses are represented with 1024 time points, and spectra with 512 complex spectral values. Given the reduced spectral resolving power of the human auditory system at high frequencies, it seems clear that without sacrificing perceptually relevant detail, HRTFs could be represented

with considerably less resolution. Use of simplified HRTFs is desirable since computational efficiency would be increased.

We have evaluated two different approaches to the problem of simplifying HRTFs. One involves smoothing the spectral representation of the HRTF, and the other attempts to model the HRTFs with a set of underlying basis functions, determined from a principal components analysis. While the smoothing technique has proven useful for display purposes, neither smoothing nor principal components analysis has yet led to significant reductions in the extent of computation required for 3-D display synthesis.

Our smoothing algorithm is based on the well-established fact that human spectral resolving power diminishes with increasing frequency. Spectral smoothing can, in general, be viewed as the convolution of the unsmoothed spectrum (magnitude and phase separately) with a filter. The extent of smoothing is given by the bandwidth of this filter, and, to some extent, its shape. In our case, the filter had a Gaussian shape and its bandwidth was set equal to the average human "critical bandwidth", the usual measure of spectral resolving power (Sharf, 1970). Since critical bandwidth increases with frequency, the resulting smoothing is greater at high frequencies. Figure 64 shows the effect of different amounts of smoothing (expressed in terms of fractions of the normal human critical bandwidth) on the HRTF measurements.

We have found that smoothing the HRTF measurements is useful for describing certain features of the HRTFs in a manner that is reasonable from a psychophysical point of view. For example, earlier in this report, estimates of the between-subjects variability in the HRTF were presented (Figures 53-56). Had smoothing not been used to remove the large, narrow peaks and troughs in the HRTFs at high frequencies, the between-subjects variability in this frequency region would have been grossly overstated. While it is possible that smoothing might be used in some way to reduce the complexity of the HRTFs used in the 3-D synthesis algorithms, we have not done so to date.

Our second attempt at simplifying HRTFs involved a principal components analysis (PCA). The central idea of PCA is to reduce the dimensionality of a data set in which there are a large number of interrelated measures, while retaining as much as possible of the variation present in the data. This reduction is accomplished by transforming the original data to a new set of measures, the principal components, which are uncorrelated. These components are extracted in an orderly fashion so that the first component reflects the majority of common variation and the remaining components reflect decreasing common variation and increasing unique variation. Thus, PCA can potentially provide a mechanism for describing the important spectral features of HRTFs and allow an HRTF for a given spatial location to be derived from a small set of "basic" functions (i.e., the principal components).

The major deterrent to using PCA to simplify HRTFs is that these functions are complex. It is inappropriate to perform traditional PCA on magnitude functions and phase functions separately and then combine the magnitude and

phase principal components to reproduce the original HRTFs. Although mathematical algorithms for complex PCA have been derived, the techniques involve sophisticated mathematics and require many hours of computer time to perform. Moreover, the computer software to implement complex PCA is not widely available. Consequently, it is difficult to evaluate the usefulness of this technique since it has been used so infrequently. Before embarking on complex PCA, we decided to investigate the usefulness of PCA for data reduction and for identifying important features of HRTFs, by first analyzing the magnitude functions.

Principal components analysis was performed on the magnitude estimates of the 144 measurements of each of the 20 subjects. Only the data in the frequency region from 200 Hz to 15000 Hz was analyzed. The general result was that approximately 90% of the variation in the 144 magnitude spectra could be accounted for by 5 principal components. That is, the frequency region between 200 and 15000 Hz could be reduced to 5 measures with very little loss of information. The amount of variation captured by 5 components did vary somewhat across subjects, ranging from 86% to 92%. Additionally, there were significant differences in the composition of the components across subjects, although there were also some similarities. This result is not surprising since we have observed large individual differences in the HRTFs above 5 kHz, and similarities below 5 kHz. The first three principal components are plotted for 5 representative subjects in Figures 65-69. The remaining two components tended to account for smaller amounts of variation in very narrow frequency regions, and thus were not included in the figures. The functions plotted in these figures reflect the contribution of each frequency to the component by it's correlation with the component. Correlations near 1.0 or -1.0 are indicative of a large amount of variation accounted for by the component (r^2 is the amount of variance explained by a component), while a correlation near 0 suggests minimal contribution to the component.

The first principal component is comprised of high correlations in the frequency region between 200 Hz and 5000 Hz for all subjects. The high correlations in this frequency region reflects the fact the 144 HRTFs are highly similar in the low frequency region. The correlations on the first component in the region between 5 kHz and 15 kHz are somewhat lower and more variable, reflecting the greater variation in the region with sound source location. Components 2 and 3 account for most of the remaining variation in the high frequencies. The pattern of correlations on Components 2 and 3 differ from subject to subject and corroborate our previous accounts of interindividual variability in the high frequency region.

In summary, principal components analysis (PCA) of the magnitude components of the HRTFs revealed that 90% of the variance could be accounted for by 5 principal components. The first of these confirmed the overall similarity of HRTF magnitude across subjects in the low frequencies, and the next two revealed large differences both across subjects and across source positions in the important 5 kHz-15 kHz region. We conclude that while principal components analysis of HRTF data in the frequency domain is useful for describing the

common features of HRTF magnitude functions, it is not appropriate for simplification or regeneration of HRTFs. Either the time-domain PCA developed by Molenaar (1985), or the complex frequency domain PCA described by Brillinger (1975) may be applicable to this problem, but, as yet, these procedures have not been thoroughly evaluated.

4.0 PSYCHOPHYSICAL EXPERIMENTS

This project involved extensive psychophysical testing. The aim of the tests was to evaluate the perceptual adequacy of auditory images synthesized from the HRTF measurements. The influence of various distortions of the HRTFs on the perceptual adequacy was of particular interest. More specifically, the psychophysical tests were intended to address the following questions:

1) What is the ability of subjects to estimate the azimuth and elevation of virtual sources produced from the subjects' own HRTF measurements, and how does this ability compare to the subjects' ability to localize real sources in free-field?

2) How is a subject's ability to localize virtual sources affected if the sources are produced from non-individualized HRTFs (e.g., from HRTFs measured on another subject or on KEMAR)?

3) What is the influence of alterations in the magnitude of the HRTF above 5 kHz on a subject's estimates of the apparent azimuth and elevation of virtual sources synthesized from the HRTF.

The psychophysical methods were identical to those described in Wightman and Kistler, 1989b. For completeness, that description is reprinted here.

4.1 Stimuli

The basic stimulus in this experiment was a train of eight 250-ms bursts of Gaussian noise (20 ms cosine-squared onset- offset ramps), with 300 ms of silence between the bursts. The noise bursts were presented at an overall level of about 70 dB SPL. The Gaussian noise was band-passed with a 10th-order digital Finite Impulse Response (FIR) band-pass filter between 200 Hz and 14 kHz. The energy spectrum of the noise was shaped (differently for each stimulus) according to an algorithm which divided the spectrum into critical bands, and assigned a random intensity (uniform distribution, 20 dB range) to the noise within each critical band. This trial-by-trial randomization of stimulus spectrum was used to prevent listeners from becoming familiar with specific stimulus or transducer characteristics.

The noise stimuli were presented either by loudspeaker or by headphones. In the former condition, the stimulus was routed to one of six small loudspeakers (Realistic Minimus-7). The loudspeakers were chosen to have similar response characteristics (+/- 5 dB from 200 Hz to 14 kHz), so no attempt was made to compensate for loudspeaker differences beyond the trial-by-trial stimulus spectral shaping described above. The loudspeakers were mounted on a semicircular steel arc, 2.76 m in diameter, the ends of which were attached to bearings directly above and below the subject's seat in an

anechoic chamber. The subject was seated on an adjustable stool such that his/her head was at the center of the arc of loudspeakers. The arc could be rotated around the vertical axis, thus allowing stimulus presentation at any azimuth, and at any one of six elevations. The loudspeakers were positioned at the following elevations relative to the horizontal plane passing through the subject's ears: 54 degrees, 36 degrees, 18 degrees, 0 degrees, -18 degrees, and -36 degrees.

For headphone conditions, the noise bursts were transduced by Sennheiser dynamic headphones (HD-340). Each headphone stimulus was digitally processed so that it would simulate a specific free-field stimulus. This processing compensated for the characteristics of the headphones, and superimposed a given subject's direction-specific outer ear characteristics (HRTF) on the stimulus (Wightman and Kistler, 1989a). Production of each stimulus involved passing a shaped burst of Gaussian noise, spectrally contoured according to the algorithm described above, through two digital filters, one for the left-ear stimulus, and the other for the right-ear stimulus. Each digital filter consisted of two cascaded sections. The first was the filter described in the companion paper (Equation 4 from Wightman and Kistler, 1989a) which includes the subject's HRTF for a given ear and source position and the inverse of the subject's headphone-to-ear-canal transfer function for that same ear. The HRTF and headphone transfer functions were measured according to the procedures described in the companion paper (Wightman and Kistler, 1989a). The second section was a zero-phase band-pass filter (200 Hz to 14 kHz) that was used to eliminate processing artifact at low and high frequencies. Finally, since the particular D/A system used to output the stimuli (Ariel DSP-16) imposed a constant 10 ms delay between left and right stimuli, a 10 ms time-shift was added to the phase response of the right band-pass filter section to compensate for the delay. Stimuli were filtered in the frequency domain, using techniques based on the "overlap and add" FFT algorithm described by Stockham (1966).

Stimuli for a given subject and a given run were precomputed (using Signal Technology Inc.'s ILS software on a DEC VAX-11/750) and stored on an IBM-PC disk. They were then converted to analog form via PC-controlled 16-bit D/A converters at a 50 kHz/channel rate. No antialiasing filters were used, since the nearest aliased components were at 36 kHz, well beyond the range of hearing. Stimuli were presented at about 70 dB SPL in free-field, and at approximately the same level under headphones. The digital processing of the headphone stimuli preserved all the interaural level and time differences, and the slight position-to-position level differences (e.g., from front to back) that existed in free-field. Figure

4 shows a block diagram of the hardware used in the psychophysical experiments.

4.2 Procedure

The aim of this experiment was to compare the apparent positions of sounds presented in free-field and under headphones. Therefore, we felt that the paradigm used to quantify apparent spatial position must be the same for both free-field and headphone listening. After considerable pilot work in which we compared the strengths and weaknesses of a number of techniques (Wightman and Kistler, 1980), we chose an "absolute judgement" technique. With this procedure, a subject indicates the apparent spatial position of a sound source by calling out numerical estimates of apparent azimuth and elevation, using standard spherical coordinates. (In our previous work with this procedure, we also asked for distance estimates.) To give some examples, a sound heard directly in front would produce a response "0,0", a sound heard on the right and slightly elevated would produce "90,10", a sound heard on the left and below the horizontal plane would produce "-90, -10", and a sound in the rear and well elevated would produce "180, 60".

We were initially concerned that our subjects would demonstrate a wide range of skill with the absolute judgement paradigm, and that this source of variance would contaminate our results. It would then be difficult to separate individual differences in localization ability from individual differences in position estimation skill. However, for several reasons, we proceeded anyway. First, our main interest was the comparison of performance in free-field with performance under headphones, and both would be measured with the absolute judgement procedure. Second, our subjects appeared to learn the procedure very quickly and produced very stable judgements. Nevertheless, all subjects were given 10 hours of experience in the free-field listening condition before final data were collected.

The free-field condition, which was tested first for 10 subjects, required subjects to estimate the apparent position of sounds delivered from 36 different positions, covering a 360 degree range of azimuths and elevations from 36 degrees below the horizontal plane to 54 degrees above it. The source locations were chosen from a list of 144 potential positions, which were those at which each subject's HRTFs had been measured (Wightman and Kistler, 1989a). The choice was made with the aim of sampling the possible range of azimuths and elevations equally. Later in the experiment, after subjects had completed testing in both free-field and headphone conditions, a second set of 36 positions was selected and 7 of the 8 subjects were tested again in both free-field and headphone conditions. Table 2 gives the coordinates of all 72 source locations,

and shows how they were divided into "low", "middle", and "high" elevations, and "front", "side", and "back" azimuths for later analysis.

At the beginning of a run in the free-field condition, subjects were blindfolded, led into the anechoic chamber, and seated at the center of the loudspeaker arc (no subject saw the inside of the anechoic chamber or the loudspeaker arrangement at any time during free-field testing). The subject was instructed to look straight ahead and not to move the head while a trial was in progress. The experimenter, who was present with the subject in the chamber in order to move the loudspeaker arc and to record the subject's responses, verified head position and stability. Each trial began with the presentation of a 15 second burst of white Gaussian noise from a loudspeaker (not one of those used for localization) mounted in front (or, in a separate condition, behind) of the subject at floor level. The purpose of this noise was to mask the sounds made by moving the loudspeaker arc, which was positioned by the experimenter during this 15 second pretrial period. When questioned later, all subjects reported that they could not detect the movement of the loudspeaker arc. After the masking noise terminated, the stimulus was presented. Recall that each stimulus consisted of eight 250 ms identical bursts of spectrally-contoured noise. During a 5 second silent period immediately after termination of the stimulus, the subject called out azimuth and elevation estimates, and the experimenter entered the responses on a data sheet (no feedback was given to the subjects). A new trial began with the experimenter repositioning the loudspeaker arc according to a script shown on the data sheet. The experimenter attempted to move the arc for about the same length of time, regardless of the required azimuth. During stimulus presentation the experimenter moved to a corner of the chamber, so as to be acoustically unobtrusive. On a given run, subjects heard a stimulus from each of the 36 locations once; the order of locations presented on each run was random. Each 36-trial run lasted about 20 minutes, and breaks of about 5 minutes were taken after each run.

The procedure for the headphone condition was nearly identical to that used for the free-field condition, except that the subjects heard the stimuli over headphones. To avoid the potential influence of visual cues, the subjects were blindfolded as in the free-field condition, even though they had seen the inside of the anechoic chamber during the acoustical measurement phase of the experiment, which came after free-field testing. They were also seated in the anechoic chamber during headphone testing. The trial sequence was the same as for the free-field condition, except that no masking noise was presented before each trial. After each stimulus was presented, and the subject called out azimuth and elevation estimates, the experimenter, who was outside the chamber and

listening over an intercom, entered the responses on a PC keyboard. As before, each run required estimates of 36 source positions, and because of the slightly faster pace, about 4 runs were completed in each 90 minute session.

Fifteen subjects participated in the psychophysical testing phase of this project. The subjects were young adults, 6 male and 9 female, with normal hearing as verified by audiometric screening at 15 dBHL. None of the subjects had any previous experience in psychoacoustical experiments, and all were naive regarding the purpose of the project. All of the subjects completed the free-field and simulated free-field conditions; only 5 of the subjects participated in all experiments described below.

Each subject completed 12 runs in the free-field condition before HRTF measurements were made. Then, 8 runs or more were completed in each of the conditions involving virtual sources. Aside from testing the free- field condition first, and the simulated free-field condition (simulations were based on each subject's own HRTF data) second, no attempt was made to present the various conditions in either a random or a counterbalanced order. Our previous work (Wightman and Kistler, 1989b) suggested that once performance had stabilized (within 5-6 runs), order effects would contribute little to the data. Analysis of the psychophysical data consisted of computing the judgement centroid (the "average" apparent direction) for each stimulus in each condition, for each subject separately. The results are presented in the form of scatterplots of actual (or intended, in the case of virtual sources) source azimuth and elevation vs. judged source azimuth and elevation (the latter given by the judgement centroid).

4.3 Free-Field versus Simulated Free-Field

Figures 70-84 show the results from the conditions in which 15 subjects localized real sources in free-field, and then localized virtual sources presented over headphones. The virtual sources, in this case, were synthesized from each subject's own HRTFs, and no distortion was introduced in the HRTFs. Table 3 summarizes the data in a form that allows assessment of both individual and average performance in various sectors of auditory space. Table 3 also presents data on intersubject differences in performance.

One important result that can be seen in the data is that in all but one case (subject SDE), subjects judge the apparent azimuth and elevation of real sources accurately. Note also that intersubject differences in localization performance appear only in the elevation components of the judgements; all subjects judge apparent source azimuth about equally well, but there are substantial differences in ability to judge apparent elevation.

The most important result is that the pattern of results from the free-field condition is duplicated with virtual sources presented over headphones. In every case, including that of the one "poor" subject, SDE, the free-field data closely match the headphone data. Where minor discrepancies appear (e.g., subject SED), they are only in the elevation component of the judgements.

These results clearly confirm the perceptual adequacy of our 3-D auditory display techniques in those conditions in which each subject's stimuli are individually tailored through the use of the subject's own HRTFs in the synthesis. The importance of individualized stimulus synthesis will be explored in the next series of experiments, using the results just discussed as a benchmark for comparison.

4.4 Use of Non-Individualized HRTFs

Two conditions were studied in which subjects localized virtual sources synthesized from HRTFs measured from other than their own ears. In one, 11 of the original subjects localized stimuli synthesized from the HRTFs measured from the ears of our "best" subject. In the other condition, 7 subjects (6 of the 11 mentioned above, plus the "best" subject) localized stimuli synthesized from KEMAR's HRTFs.

Figures 85-95 show the data obtained from the first condition. The results can be easily summarized. First, the use of another subject's HRTFs for stimulus synthesis causes no more than minor alterations in the azimuth components of judgements of apparent source position. Second, with regard to the elevation components of the judgements, performance is never as good as with stimuli synthesized from the subject's own HRTFs. In other words, if a "good" subject (i.e., a subject who judges apparent elevation accurately) listens to stimuli synthesized from another "good" subject's HRTFs, there are slight degradations of elevation performance. If a "poor" subject (e.g., SDE) listens to stimuli synthesized from a "good" subject, there is no improvement. Table 4 summarizes the data from this condition.

Figures 96-102 show data from the second condition, in which seven subjects localized stimuli synthesized from HRTF measurements made on KEMAR. The results are similar to those from the first condition. First, judgements of apparent source azimuth are relatively unaffected by the use of KEMAR's HRTFs. However, there was a substantial increase in front/back confusions for most subjects. Second, for about half of the subjects, apparent source elevation was only slightly distorted, and for the other half it was badly distorted. These distortions, which are manifest by lowered apparent elevation for sources at positive elevations, occurred with sources at all azimuths. An example of this effect can be seen in Figure 98, which shows that for all sources (regardless of azimuth) with elevations above zero, judged elevation is close to zero. It is tempting to attribute the differences between the subjects' performance with KEMAR-based stimuli to HRTF differences. However, we have been unable to trace any subject's poor (or good) performance with KEMAR-based stimuli to differences (or similarities) between that subject's own HRTFs and KEMAR's. Table 5 presents the summary data from this condition.

4.5 Effects of High Frequency Distortions in the HRTF

Two aspects of the results reported here and elsewhere (e.g., Wightman and Kistler, 1989b) argue indirectly for the importance of high frequencies to the perception of source elevation. First, analysis of the HRTFs reveals that between-subjects differences in the HRTFs are greatest at high frequencies. Second, results from the experiments with nonindividualized HRTFs suggest that accurate elevation perception requires that a subject's own HRTF be used to synthesize the stimuli. Use of a different subject's HRTF would, presumably, distort the high frequency region more than the low frequency region.

The experiment described here was designed to assess the importance of various frequency regions directly. Seven of the original 15 subjects localized virtual sources (synthesized from their own HRTFs) in which the energy in various frequency regions had been removed by filtering. The filtering was accomplished digitally, with a high-order FIR filter, so stop-band attenuation was at least 80 dB, and the transition band was very narrow (filter skirts were very steep). Four conditions were studied: 1) 5 kHz low-pass (all energy above 5 kHz removed); 2) 5 kHz high-pass (all energy below 5 kHz removed); 3) 10 kHz low-pass, and 4) 10 kHz high-pass. Six of the seven subjects were tested in all four conditions; the remaining subject was tested in three of the four conditions. The psychophysical procedure was identical to that used in the other experiments described above. Since performance was comparable across all subjects, only the data from one subject (who completed all four conditions) will be shown. Figures 103-106 show the results from the filtering experiment in scatterplot form, and Tables 6-9 present the summary data for conditions 1-4, respectively.

Note first, that in all conditions except the 10 kHz high-pass condition, apparent source azimuth was virtually unaffected by filtering. Apparent elevation, however, was dramatically affected in some conditions. Consider the 5 kHz low-pass condition (Figure 103). Here, with all energy above 5 kHz removed, the apparent elevation of all sources is close to zero degrees. This strongly suggests that the cues for source elevation are encoded in frequencies above 5 kHz. The results from the 5 kHz high-pass condition (Figure 104), in which elevation perception appears normal, confirm the importance of energy above 5 kHz. Next, consider the 10 kHz high-pass condition (Figure 105). With all low-frequency information removed, elevation perception is also seriously degraded. The results from the former three conditions, combined with the results from the 10 kHz low-pass condition (Figure 106), which show normal elevation perception, indicate that the major cues to source elevation lie in the 5-10 kHz frequency region.

5.0 IMPLICATIONS FOR PRACTICAL APPLICATION OF 3-D AUDITORY DISPLAY TECHNOLOGY

The psychophysical results lead us to conclude that for the most veridical 3-D auditory display, a listener's own HRTFs should be used to synthesize the virtual sound sources. If only azimuth information is to be conveyed in the display, the synthesis requirements can be relaxed considerably. However, if the apparent elevation of a source is important, individualized HRTFs appear to be essential. This conclusion is quite different from that reached by Butler and Belendiuk (1977) in their frequently-cited paper on median-plane localization. In the Butler and Belendiuk study, four listeners localized noises that had been recorded from microphones placed either in their own ears or in the ears of other listeners. While three of the four subjects showed no effect or degradation in localization performance with non-individualized stimuli, one of the subjects appeared to localize more proficiently with stimuli recorded from one of the other subject's ears than from his/her own. However, we feel this result must be interpreted with great caution, for several reasons. First, the performance of the one unusual subject was generally quite poor, and, in fact, was at chance in free-field. Second, only one of the four subjects showed the effect. Third, the task involved not localization, but identification of a target source from a small group (5) of sources, arranged only on the median plane. The generalizability of these results to localization conditions such as we have studied seems questionable. It is possible (and we feel quite likely) that, over time, with visual and kinesthetic feedback (neither of which was available in our study), listeners could become quite proficient localizing stimuli synthesized from non-individualized HRTF data (Searle, 1982).

The data also suggest that the importance of the 5-10 kHz frequency region to elevation perception (revealed best by the filtering experiments) must be recognized, if veridical elevation perception is expected. Great care must be taken in making HRTF measurements and in synthesizing stimuli to preserve spectral information in this region. This means that: 1) the probe microphone used to measure the HRTF must be sufficiently close to the eardrum to avoid a standing wave null in the 5- 10 kHz region; and 2) signal/noise ratio in the 5-10 kHz region must be sufficiently high to preserve spectral detail; only high-energy stimuli (i.e., not transients) and low-noise environments (i.e., sound-proofed, anechoic rooms) should be used for HRTF measurements.

It is quite possible that future research will lead to stimulus synthesis techniques that do not depend on individualized HRTF measurements. For example, it may be the case that mathematical models of the external ear, based on a small number of anatomical measurements, could serve as the basis for the synthesis algorithms. However, at this time, it is our conclusion that only carefully measured individualized HRTF measurements can convey completely accurate source position information for a 3-D auditory display.

6.0 REFERENCES

Butler, R. A. and Belendiuk, K. (1977). "Spectral cues utilized in the localization of sound in the median sagittal plane," J. Acoust. Soc. Am., 61, 1264-1269.

Brillinger, D. R. (1975). Time Series: Data Analysis and Theory, (Holt, Rinehart, and Winston, New York).

Chan, J. and Geisler, C. D. (1989). "Estimation of eardrum acoustic pressure from remote points in the ear canal," J. Acoust. Soc. Am., In Press.

Mehrgardt, S. and Mellert, V. (1977). "Transformation characteristics of the external human ear," J. Acoust. Soc. Am., 61, 1567-1576.

Molenaar, P. C. M. (1985). "A dynamic factor model for the analysis of multivariate time series," Psychometrika, 50, 181-202.

Searle, C. L. (1982). "A model of auditory localization: Peripheral Constraints," in Localization of Sound: Theory and Applications, edited by R. W. Gatehouse (Amphora Press, Groton, Conn.).

Sharf, B. (1970). "Critical bands," in Foundations of Modern Auditory Theory - I, edited by J. Tobias (Academic Press, New York).

Shaw, E. A. G. (1980). "The acoustics of the external ear", in Acoustical Factors Affecting Hearing Aid Performance, edited by G. A. Studebaker and I. Hochberg (University Park Press, Baltimore).

Shaw, E. A. G. (1965). "Earcanal pressure generated by a free sound field," J. Acoust. Soc. Am., 39, 465-470.

Schroeder, M. R. (1970). "Synthesis of low-peak-factor signals and binary sequences with low autocorrelation," IEEE Trans. Inform. Theory, 16, 85-89.

Stockham, T. G. (1966). "High speed convolution and correlation," Proc. AFIPS Spring Joint Comp. Conf., 28, 229-233.

Wightman, F. L. and Kistler, D. J. (1980). "A new "look" at auditory space perception," in Psychophysical, Physiological, and Behavioral Studies in Hearing, edited by G. van den Brink and F. Bilsen (University Press, Delft).

Wightman, F. L. and Kistler, D. J. (1989a). "Headphone simulation of free-field listening. I: Stimulus synthesis," J. Acoust. Soc. Am., 85, 858-867.

Wightman, F. L. and Kistler, D. J. (1989b). "Headphone simulation of free-field listening - II: Psychophysical validation," J. Acoust. Soc. Am., 85, 868-878.

Table 1

Estimates of Canal Length
and Distance from Eardrum

ID	Ear	Visual Canal Length	Acoustic Canal Length	Visual Eardrum Distance	Acoustic Eardrum Distance
SDE	L	20.5	23.6	4.5	7.6
	R	21.0	23.2	6.0	8.2
SDL	L	25.0	24.9	10.8	10.7
	R	23.0	23.9	8.5	9.4
SDO	L	25.0	25.6	6.5	7.1
	R	24.0	25.5	7.0	8.5
SDP	L	27.0	27.0	9.5	9.5
	R	26.5	28.9	6.5	8.9
SED	L	22.0	23.5	7.7	9.2
	R	22.0	23.4	6.8	8.2
SER	L	24.5	24.2	12.1	11.8
	R	20.5	24.2	7.5	11.2
SET	L	20.0	20.1	5.9	6.0
	R	19.0	19.7	4.9	5.6
SGB	L	24.5	26.2	10.2	11.9
	R	26.0	26.1	14.6	14.7
SGD	L	20.0	21.1	4.0	5.1
	R	22.0	21.8	5.0	4.8
SGE	L	19.5	22.0	3.9	6.4
	R	20.5	22.8	3.7	6.0
SGG	L	22.5	19.7	10.9	8.1
	R	20.0	20.5	8.8	9.3
SHD	L	19.5	<20.3	3.9	<4.7
	R	18.0	20.2	3.0	5.2
SHF	L	26.0	29.2	3.5	6.7
	R	25.0	28.1	2.5	5.6

Table 2

Source Positions

	Low		Middle		High	
	Azim.	Elev.	Azim.	Elev.	Azim.	Elev.
Front						
	-15	-36	-15	0	-45	36
	-45	-36	-45	0	0	36
	30	-36	0	0	15	36
	45	-36	15	0	30	36
	-15	-18	45	0	-30	54
	0	-18	-15	18	-45	54
	30	-18	-30	18	15	54
	45	-18	-45	18		
			0	18		
			45	18		
Side						
	90	-36	-75	0	-60	36
	105	-36	-90	0	120	36
	-60	-18	-105	0	90	36
	-90	-18	60	0	-105	54
	75	-18	90	0	75	54
	105	-18	-75	18	90	54
			-90	18		
			-120	18		
			75	18		
			105	18		
Back						
	-135	-36	-135	0	-135	36
	-150	-36	150	0	-150	36
	135	-36	165	0	165	36
	150	-36	180	0	180	36
	180	-36	135	18	-150	54
	-135	-18	150	18	-165	54
	-150	-18	165	18	180	54
	-165	-18				
	135	-18				

TABLE 3

Regional Measures of Performance in Free-Field and Simulated Localization

		Low elevations			Middle elevations			High elevations		
Azimuth	HD	Angle of Error	κ^{-1}	% reversals	Angle of Error	κ^{-1}	% reversals	Angle of error	κ^{-1}	% reversals
Front	SDE	37.4 (34.9)	.08 (.05)	21 (46)	16.2 (19.7)	.04 (.05)	4 (21)	28.3 (40.0)	.10 (.10)	19 (29)
	SDH	20.0 (16.0)	.04 (.03)	4 (17)	21.1 (37.4)	.03 (.08)	1 (8)	19.3 (26.2)	.03 (.05)	12 (58)
	SDL	21.9 (22.2)	.06 (.03)	3 (0)	25.9 (29.1)	.07 (.07)	1 (11)	27.8 (31.3)	.07 (.14)	44 (81)
	SDM	19.4 (23.1)	.05 (.05)	0 (3)	16.3 (15.3)	.02 (.01)	0 (5)	24.0 (21.9)	.01 (.10)	4 (20)
	SDO	14.2 (15.2)	.05 (.02)	1 (0)	14.5 (21.8)	.04 (.06)	0 (0)	25.7 (34.8)	.04 (.15)	7 (39)
	SDP	10.9 (22.3)	.02 (.06)	0 (6)	9.4 (13.9)	.01 (.03)	0 (9)	17.1 (23.2)	.05 (.06)	5 (34)
	SED	19.2 (19.0)	.07 (.06)	0 (2)	21.2 (15.1)	.05 (.03)	0 (0)	31.6 (29.1)	.08 (.10)	19 (30)
	SER	19.6 (16.7)	.08 (.03)	1 (0)	17.9 (16.7)	.05 (.04)	2 (0)	27.4 (24.5)	.07 (.15)	11 (2)
	SET	16.8 (20.4)	.06 (.08)	1 (2)	24.3 (25.3)	.08 (.09)	8 (7)	38.8 (32.2)	.03 (.14)	24 (33)
	SGB	21.5 (28.2)	.07 (.13)	2 (41)	35.7 (21.2)	.10 (.08)	1 (54)	34.9 (25.7)	.02 (.11)	14 (68)
	SGD	12.5 (16.0)	.05 (.04)	6 (20)	15.5 (16.7)	.05 (.04)	0 (5)	13.8 (31.5)	.05 (.10)	5 (58)
	SGE	19.6 (24.8)	.06 (.03)	0 (0)	20.4 (29.0)	.02 (.06)	2 (5)	29.1 (23.4)	.15 (.11)	46 (43)
	SGG	27.6 (16.2)	.07 (.03)	0 (2)	22.4 (16.0)	.08 (.03)	4 (4)	35.3 (20.8)	.18 (.05)	17 (20)
	SHD	17.5 (17.5)	.06 (.05)	3 (0)	39.3 (23.5)	.13 (.06)	8 (11)	37.9 (27.3)	.11 (.15)	9 (44)
	SHF	19.4 (15.8)	.09 (.04)	0 (4)	18.5 (18.5)	.07 (.05)	1 (4)	27.8 (43.0)	.12 (.15)	51 (50)
	Mean	20.2 (20.6)	.06 (.05)	3 (10)	21.2 (21.3)	.06 (.05)	2 (10)	27.9 (29.0)	.07 (.11)	19 (41)
Side	SDE	22.9 (33.4)	.03 (.06)	6 (14)	17.0 (18.3)	.04 (.05)	19 (20)	24.5 (37.1)	.10 (.15)	28 (28)
	SDH	14.7 (10.4)	.03 (.02)	17 (28)	14.3 (13.8)	.03 (.02)	6 (3)	17.9 (16.0)	.02 (.03)	25 (11)
	SDL	18.4 (21.3)	.04 (.04)	0 (1)	18.7 (14.2)	.04 (.02)	9 (6)	24.7 (29.5)	.09 (.11)	16 (32)
	SDM	19.4 (15.6)	.07 (.04)	11 (10)	17.9 (19.7)	.07 (.05)	7 (12)	19.7 (26.8)	.05 (.15)	21 (10)
	SDO	13.2 (14.3)	.02 (.03)	7 (21)	16.6 (17.4)	.05 (.03)	5 (12)	21.7 (27.4)	.04 (.12)	14 (16)
	SDP	13.8 (15.1)	.06 (.04)	4 (0)	15.0 (10.1)	.02 (.02)	8 (6)	17.2 (16.9)	.05 (.03)	7 (4)
	SED	20.0 (17.1)	.05 (.03)	2 (7)	16.5 (12.8)	.05 (.03)	2 (0)	27.5 (37.9)	.14 (.09)	10 (17)
	SER	20.2 (18.5)	.07 (.07)	2 (9)	13.9 (17.3)	.04 (.04)	7 (7)	26.7 (22.4)	.08 (.06)	22 (15)
	SET	18.6 (16.4)	.07 (.05)	9 (17)	14.5 (16.8)	.03 (.07)	7 (9)	23.9 (33.2)	.07 (.24)	11 (26)
	SGB	21.7 (20.8)	.09 (.04)	6 (12)	20.3 (16.5)	.07 (.03)	17 (28)	27.2 (28.3)	.09 (.10)	14 (22)
	SGD	20.1 (16.4)	.10 (.04)	4 (5)	15.8 (14.0)	.04 (.03)	0 (4)	23.9 (32.1)	.08 (.07)	6 (17)
	SGE	23.8 (19.9)	.10 (.05)	2 (13)	16.9 (17.7)	.06 (.04)	21 (14)	21.5 (16.1)	.07 (.05)	27 (38)
	SGG	23.6 (14.0)	.05 (.03)	18 (1)	13.6 (10.2)	.03 (.02)	14 (4)	31.5 (19.8)	.17 (.04)	20 (8)
	SHD	17.9 (18.8)	.05 (.04)	12 (7)	19.3 (16.5)	.09 (.03)	12 (14)	20.1 (23.2)	.07 (.09)	16 (31)
	SHF	19.6 (23.4)	.06 (.11)	6 (11)	17.1 (13.3)	.05 (.08)	10 (13)	26.6 (22.7)	.11 (.07)	22 (17)
	Mean	19.2 (18.4)	.06 (.05)	7 (10)	16.5 (15.2)	.05 (.01)	10 (10)	23.6 (26.0)	.08 (.09)	17 (19)
Back	SDE	19.6 (17.6)	.04 (.04)	4 (1)	22.9 (24.2)	.08 (.07)	4 (2)	44.8 (49.3)	.08 (.09)	11 (21)
	SDH	22.4 (20.2)	.03 (.02)	0 (2)	27.3 (25.9)	.04 (.03)	0 (0)	24.8 (27.1)	.05 (.05)	0 (0)
	SDL	21.9 (21.5)	.04 (.03)	0 (2)	14.5 (17.7)	.04 (.03)	0 (0)	24.6 (27.7)	.12 (.15)	2 (10)
	SDM	16.0 (16.2)	.02 (.06)	0 (8)	18.4 (20.9)	.05 (.07)	0 (0)	18.4 (31.9)	.08 (.21)	0 (15)
	SDO	20.8 (16.0)	.02 (.02)	0 (0)	28.9 (22.3)	.04 (.10)	0 (2)	35.1 (33.3)	.08 (.11)	5 (15)
	SDP	15.7 (11.7)	.02 (.01)	0 (0)	14.0 (11.9)	.02 (.01)	1 (0)	21.3 (18.7)	.06 (.04)	0 (0)
	SED	23.2 (22.5)	.03 (.02)	1 (0)	20.5 (14.1)	.08 (.04)	0 (4)	31.3 (32.9)	.16 (.06)	3 (5)
	SER	13.9 (15.8)	.03 (.03)	0 (1)	20.5 (21.3)	.05 (.08)	0 (9)	33.0 (29.2)	.04 (.10)	3 (41)
	SET	18.5 (18.8)	.05 (.06)	0 (8)	21.8 (20.7)	.09 (.09)	0 (2)	34.6 (38.0)	.06 (.27)	2 (39)
	SGB	20.4 (18.0)	.09 (.07)	0 (0)	22.1 (17.8)	.09 (.05)	0 (0)	25.2 (25.6)	.05 (.13)	0 (0)
	SGD	18.0 (14.5)	.07 (.04)	0 (1)	27.6 (25.0)	.10 (.07)	13 (30)	23.2 (31.1)	.09 (.13)	2 (19)
	SGE	26.0 (19.2)	.06 (.03)	7 (24)	15.6 (16.0)	.05 (.02)	4 (2)	19.2 (24.2)	.08 (.11)	4 (35)
	SGG	22.6 (13.1)	.05 (.01)	3 (0)	19.6 (14.6)	.06 (.03)	1 (2)	40.6 (24.8)	.08 (.05)	3 (4)
	SHD	14.5 (21.1)	.05 (.08)	0 (4)	19.6 (17.4)	.08 (.05)	1 (0)	34.2 (32.0)	.15 (.17)	9 (26)
	SHF	12.1 (15.2)	.03 (.05)	2 (7)	15.9 (20.8)	.06 (.07)	0 (0)	27.0 (38.2)	.12 (.13)	6 (29)
	Mean	19.0 (17.4)	.04 (.04)	1 (4)	20.6 (19.4)	.06 (.05)	2 (4)	29.2 (30.9)	.09 (.12)	4 (17)

TABLE 4

Regional Measures of Performance with a "Good" Subject's HRTFs

		Low elevations			Middle elevations			High elevations		
Azimuth	ID	Angle of Error	κ^{-1}	% reversals	Angle of Error	κ^{-1}	% reversals	Angle of error	κ^{-1}	% reversals
Front	SDE	34.9	.11	53	18.9	.07	34	45.8	.09	75
	SDH	22.7	.05	50	31.5	.11	21	30.8	.07	75
	SDL	28.2	.05	29	30.8	.04	13	24.1	.10	75
	SDP	40.1	.05	65	36.2	.09	50	31.1	.14	65
	SED	20.5	.05	0	11.7	.03	0	32.0	.13	43
	SER	20.6	.05	5	21.6	.05	4	34.4	.20	29
	SET	20.4	.09	2	30.7	.14	19	36.4	.21	68
	SGB	29.6	.12	16	30.5	.13	23	35.5	.20	69
	SGE	27.1	.02	0	28.7	.06	14	30.1	.15	58
	SGG	18.6	.07	16	11.9	.02	25	29.6	.05	56
	SHD	18.2	.04	0	20.5	.08	8	36.8	.20	55
	Mean	25.5	.06	21	24.8	.07	19.2	33.3	.14	60.7
Side	SDE	31.1	.05	8	18.6	.05	10	35.8	.20	25
	SDH	20.5	.02	33	15.2	.03	13	20.3	.06	13
	SDL	32.7	.08	33	15.4	.03	0	28.4	.17	29
	SDP	33.2	.03	33	22.1	.05	10	34.1	.13	35
	SED	23.7	.02	9	15.1	.02	6	32.8	.09	9
	SER	17.3	.04	8	26.8	.07	8	37.0	.28	31
	SET	18.5	.03	10	15.0	.03	10	26.2	.12	35
	SGB	23.5	.09	13	25.5	.11	8	30.9	.17	36
	SGE	23.1	.02	25	24.0	.04	10	27.9	.05	46
	SGG	17.5	.04	33	11.6	.02	15	26.0	.04	22
	SHD	17.3	.04	13	15.7	.06	20	20.5	.08	13
	Mean	23.5	.04	20	18.6	.05	10	29.1	.13	27
Back	SDE	14.3	.04	10	21.9	.07	13	53.2	.07	3
	SDH	28.5	.06	7	27.8	.08	0	24.8	.08	4
	SDL	28.9	.03	10	21.6	.07	0	26.0	.14	13
	SDP	24.1	.02	0	14.1	.04	25	34.8	.07	0
	SED	27.4	.04	4	22.0	.08	5	39.5	.28	17
	SER	20.3	.03	7	27.0	.15	41	38.6	.25	27
	SET	20.1	.06	4	25.1	.11	13	35.5	.28	11
	SGB	18.7	.05	0	22.1	.10	2	37.5	.27	7
	SGE	22.3	.02	0	17.3	.06	6	33.1	.19	44
	SGG	12.9	.02	3	11 3	.02	0	22.3	.05	6
	SHD	23.4	.04	0	25.0	.07	0	31.4	.15	3
	Mean	21.9	.04	4	21.4	.08	10	34.3	.17	12

TABLE 5

Regional Measures of Performance with KEMAR's HRTFs

		Low elevations			Middle elevations			High elevations		
Azimuth	ID	Angle of Error	κ^{-1}	% reversals	Angle of Error	κ^{-1}	% reversals	Angle of error	κ^{-1}	% reversals
Front	SDL	27.9	.05	69	18.9	.05	91	31.6	.03	100
	SDO	26.1	.02	0	21.3	.07	9	29.7	.16	9
	SET	22.2	.03	59	22.5	.09	75	43.4	.06	100
	SGB	33.4	.03	100	16.6	.05	97	39.9	.03	100
	SGE	31.4	.05	44	23.6	.10	50	32.7	.07	91
	SGG	16.4	.08	28	14.0	.04	25	27.7	.04	53
	SHD	24.1	.09	44	19.5	.06	44	36.6	.07	79
	Mean	25.9	.05	49	19.5	.07	56	34.5	.07	76
Side	SDL	26.5	.07	17	26.3	.06	8	30.9	.05	25
	SDO	19.5	.04	17	18.6	.04	15	25.2	.10	34
	SET	20.9	.03	8	13.9	.02	18	50.5	.06	28
	SGB	35.5	.09	33	17.1	.03	20	49.0	.12	25
	SGE	19.1	.03	8	22.1	.04	9	36.9	.09	30
	SGG	23.2	.07	0	10.1	.01	10	26.3	.04	19
	SHD	21.5	.05	3	16.9	.03	3	26.3	.07	6
	Mean	23.7	.05	12	17.9	.03	12	35.0	.08	24
Back	SDL	31.6	.02	0	16.7	.04	4	32.9	.07	0
	SDO	30.5	.04	20	23.9	.04	96	30.2	.20	50
	SET	24.7	.04	0	29.2	.05	21	42.6	.10	3
	SGB	27.3	.03	0	15.6	.05	0	46.6	.09	0
	SGE	25.9	.03	1	17.3	.04	0	35.6	.10	11
	SGG	13.4	.02	0	12.9	.02	0	21.7	.05	9
	SHD	22.5	.03	0	22.7	.07	0	33.0	.06	0
	Mean	25.1	.03	3	19.8	.04	17.3	34.7	.10	10

TABLE 6

Regional Measures of Performance with 5 kHz Low-Pass Stimuli

		Low elevations			Middle elevations			High elevations		
Azimuth	ID	Angle of Error	κ^{-1}	% reversals	Angle of Error	κ^{-1}	% reversals	Angle of error	κ^{-1}	% reversals
Front	SDH	30.3	.07	78	25.5	.05	10	23.4	.06	50
	SDL	28.2	.05	9	15.6	.05	8	44.7	.04	38
	SDO	18.9	.02	13	20.2	.04	4	53.2	.07	8
	SER	49.3	.08	16	29.4	.06	10	16.2	.04	21
	SET	44.1	.07	6	22.8	.07	2	30.1	.10	13
	SGB	35.2	.09	69	21.0	.05	54	41.3	.06	75
	SGE	76.1	.14	16	45.4	.08	19	25.0	.12	25
	Mean	40.3	.07	30	25.7	.06	15	33.4	.07	33
Side	SDH	14.6	.05	28	21.8	.03	31	35.9	.02	50
	SDL	21.4	.03	19	19.8	.06	21	44.7	.10	0
	SDO	15.2	.03	22	19.3	.02	17	60.3	.08	6
	SER	38.3	.04	25	20.8	.09	29	27.6	.10	0
	SET	29.6	.03	25	17.0	.04	48	30.2	.07	0
	SGB	30.8	.03	25	18.7	.03	23	38.6	.11	31
	SGE	26.7	.02	25	17.3	.04	19	23.1	.08	19
	Mean	25.2	.03	24	19.2	.04	27	37.2	.08	15
Back	SDH	19.8	.02	0	32.6	.04	6	45.2	.09	17
	SDL	26.3	.06	19	13.9	.03	22	46.5	.05	21
	SDO	16.5	.04	41	19.6	.04	59	53.7	.01	100
	SER	26.6	.10	13	21.8	.06	13	37.2	.20	29
	SET	27.3	.06	38	20.2	.07	44	34.0	.09	79
	SGB	19.7	.04	0	18.2	.07	3	42.6	.05	4
	SGE	24.9	.08	22	25.7	.10	16	31.1	.20	54
	Mean	23.0	.06	19	21.7	.06	23	41.5	.10	43

TABLE 7

Regional Measures of Performance with 5 kHz High-Pass Stimuli

		Low elevations			Middle elevations			High elevations		
Azimuth	ID	Angle of Error	κ^{-1}	% reversals	Angle of Error	κ^{-1}	% reversals	Angle of error	κ^{-1}	% reversals
Front	SDH	22.3	.06	78	27.5	.07	88	20.1	.06	100
	SDL	38.0	.17	66	19.7	.08	75	28.4	.10	100
	SDO	20.0	.03	6	21.1	.05	34	28.1	.13	72
	SER	26.2	.05	6	22.6	.04	3	28.2	.12	69
	SET	25.0	.08	16	39.4	.14	28	28.4	.12	84
	SGB	30.2	.10	60	17.3	.04	80	36.3	.04	100
	SGE	29.0	.08	31	22.8	.08	65	38.4	.10	88
	Mean	27.2	.08	38	24.3	.07	53	29.7	.10	88
Side	SDH	10.6	.02	33	14.1	.03	28	28.3	.04	25
	SDL	34.8	.08	25	20.7	.07	33	43.9	.08	22
	SDO	13.8	.02	25	30.6	.03	40	30.3	.08	25
	SER	19.5	.04	17	14.9	.02	30	28.1	.09	9
	SET	16.3	.03	29	15.4	.04	23	33.9	.12	25
	SGB	29.1	.04	30	36.2	.11	36	53.3	.09	25
	SGE	21.2	.05	18	25.8	.05	37	37.7	.10	27
	Mean	20.8	.04	25	22.5	.05	32	36.5	.09	23
Back	SDH	22.7	.02	0	30.5	.09	0	27.6	.08	0
	SDL	26.4	.06	0	19.3	.05	0	33.0	.11	6
	SDO	22.9	.05	0	25.4	.07	17	22.2	.07	3
	SER	20.0	.04	0	19.0	.08	0	18.3	.07	9
	SET	25.0	.11	8	25.4	.10	13	32.6	.16	9
	SGB	21.5	.06	0	23.1	.06	0	38.6	.04	0
	SGE	23.8	.05	14	18.2	.01	0	42.3	.12	15
	Mean	23.3	.06	4	21.7	.06	5	31.2	.10	7

TABLE 8

Regional Measures of Performance with 10 kHz Low-Pass Stimuli

		Low elevations			Middle elevations			High elevations		
Azimuth	ID	Angle of Error	κ^{-1}	% reversals	Angle of Error	κ^{-1}	% reversals	Angle of error	κ^{-1}	% reversals
Front	SDL	47.3	.09	38	39.8	.06	25	16.6	.05	92
	SDO	38.3	.34	13	34.3	.09	2	28.5	.04	38
	SER	38.9	.26	53	35.8	.09	27	27.4	.03	33
	SET	23.3	.05	20	22.7	.10	15	33.2	.19	60
	SGB	44.2	.16	88	32.8	.09	85	29.7	.13	92
	SGE	48.8	.40	46	54.3	.12	24	36.5	.03	52
	Mean	40.1	.22	43	36.6	.09	30	28.7	.08	61
Side	SDL	23.5	.05	16	22.6	.06	8	16.7	.06	19
	SDO	22.0	.04	28	28.8	.08	21	17.0	.03	31
	SER	16.2	.04	3	22.9	.04	27	22.0	.03	25
	SET	18.6	.04	20	17.5	.03	28	36.5	.10	40
	SGB	22.6	.06	34	19.6	.05	17	41.7	.15	50
	SGE	19.2	.04	7	20.8	.05	5	17.2	.02	36
	Mean	20.4	.05	18	22.7	.05	18	25.2	.07	34
Back	SDL	28.0	.04	0	18.8	.07	0	23.0	.07	4
	SDO	13.4	.01	0	35.9	.10	13	26.6	.08	4
	SER	16.0	.03	0	28.2	.10	0	26.7	.07	13
	SET	23.8	.08	30	22.2	.08	13	28.8	.16	3
	SGB	19.2	.08	0	21.6	.12	3	19.2	.09	0
	SGE	12.8	.03	4	32.1	.13	0	38.5	.01	5
	Mean	18.9	.05	6	26.5	.10	5	27.1	.08	5

TABLE 9

Regional Measures of Performance on the 10 kHz High-Pass Stimuli

		Low elevations			Middle elevations			High elevations		
Azimuth	ID	Angle of Error	κ^{-1}	% reversals	Angle of Error	κ^{-1}	% reversals	Angle of error	κ^{-1}	% reversals
Front	SDH	17.7	.02	100	40.5	.03	100	73.3	.03	100
	SDL	28.4	.04	91	19.0	.04	97	49.5	.04	100
	SDO	26.1	.07	41	23.9	.04	72	53.0	.06	100
	SER	27.6	.11	79	30.6	.12	79	52.7	.21	83
	SET	34.7	.13	84	39.9	.09	91	66.1	.09	97
	SGB	27.5	.05	100	24.7	.04	100	58.3	.03	100
	SGE	28.0	.05	72	25.5	.04	75	52.8	.05	97
	Mean	27.1	.07	81	29.2	.06	88	57.8	.07	97
Side	SDH	21.1	.01	33	35.2	.01	38	78.6	.01	25
	SDL	43.1	.07	33	33.6	.03	40	57.3	.04	22
	SDO	25.5	.03	33	46.2	.02	40	69.4	.03	25
	SER	28.8	.04	28	39.4	.13	37	62.8	.18	25
	SET	30.8	.19	21	25.0	.06	28	79.8	.10	19
	SGB	39.5	.06	33	47.4	.08	40	66.3	.09	25
	SGE	33.6	.08	29	39.4	.05	38	65.8	.04	28
	Mean	31.8	.07	30	38.0	.05	37	68.6	.07	24
Back	SDH	25.4	.02	0	44.7	.02	0	77.7	.08	0
	SDL	30.8	.04	0	16.9	.03	0	51.5	.07	0
	SDO	19.1	.01	0	20.0	.05	0	55.6	10	6
	SER	20.2	.04	0	34.9	.14	0	49.5	.15	4
	SET	30.3	.04	3	37.7	.05	0	66.7	.05	0
	SGB	25.5	.08	3	25.1	.03	0	59.4	02	0
	SGE	24.9	.04	3	21.7	.02	0	55.0	.05	13
	Mean	25.2	.01	1	28.7	.05	0	59.3	.07	3

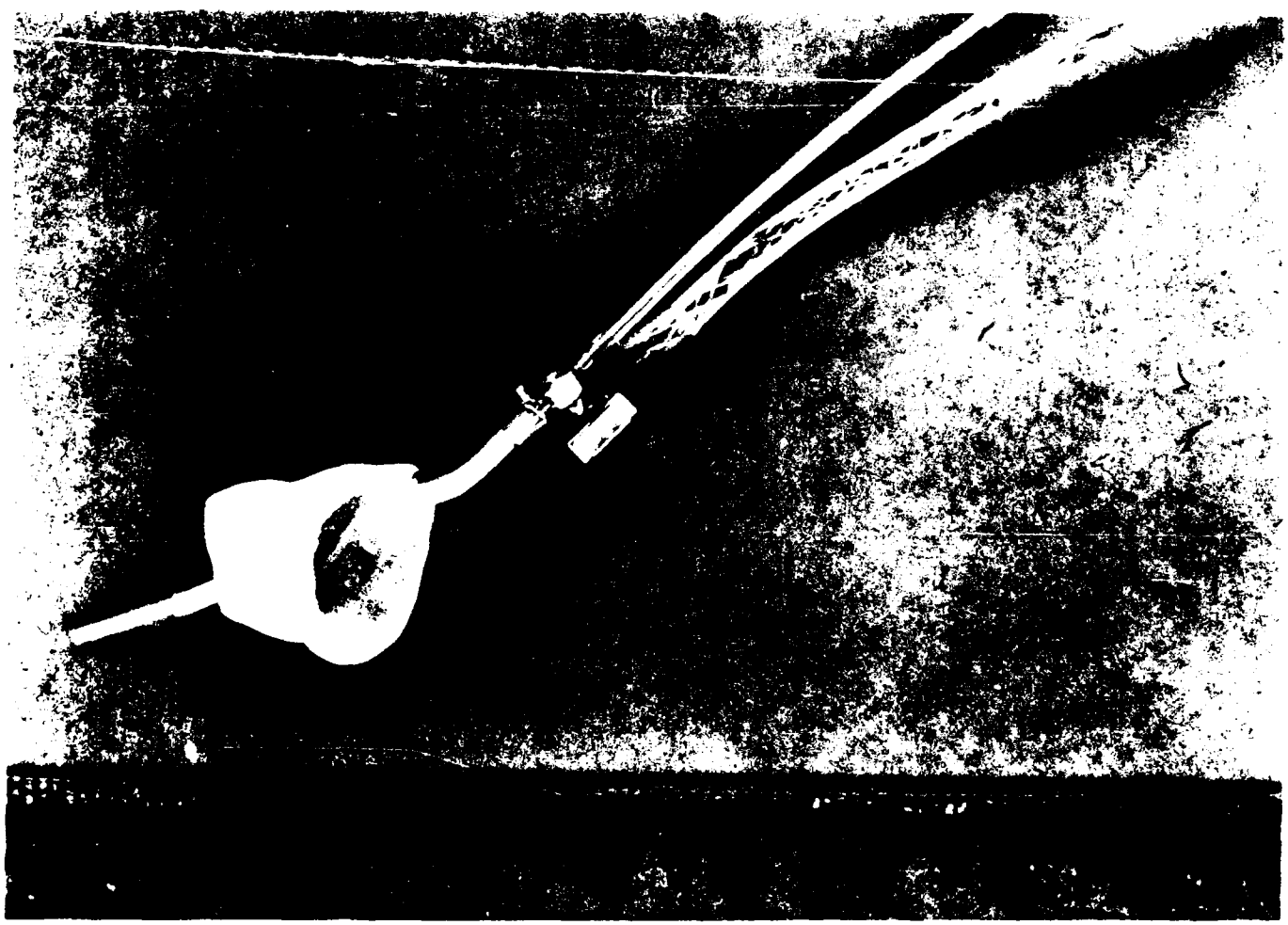


Figure 1. Photograph of one of the two ETYMOTIC microphones used to measure HRTFs from inside subjects' ear canals. The thin silicone probe tube is less than 1 mm in diameter

Figure 2. Photograph of a custom lucite earmold assembly, with the probe microphone in place, that was used in subjects' ear canalsto measure HRTFs. Note that the earmold is trimmed and boredout so that its acoustical effects would be minimal

Figure 3. Photograph of the inside of the anechoic chamber used both for HRTF measurements and for psychophysical testing. During psychophysical testing, the subject is blindfolded and the loudspeaker arc is moved by an assistant. During HRTFmeasurements, the subject moves the loudspeaker arc

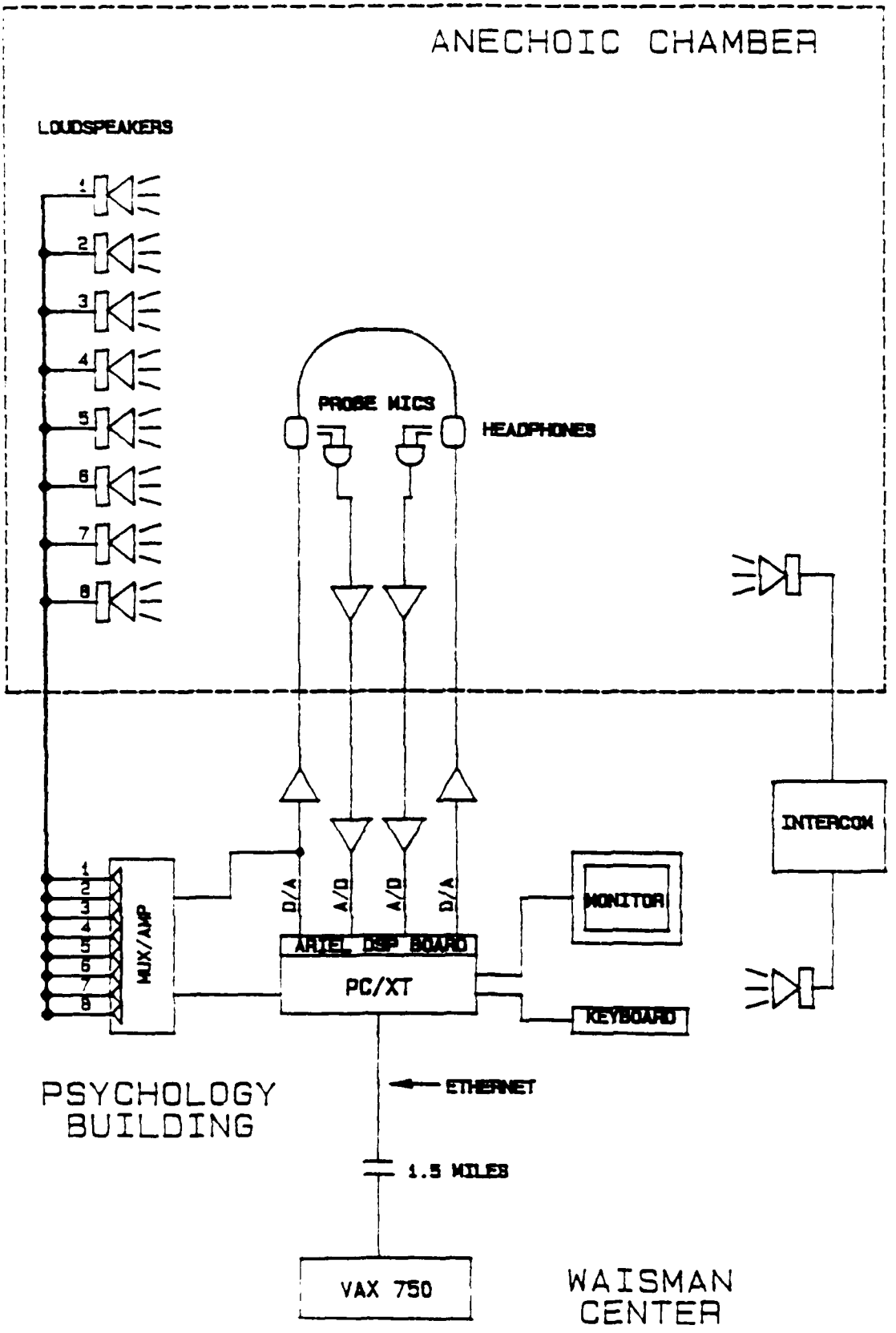


Figure 4. Block diagram showing the major hardware components in the set-up used both for HRTF measurements and for psychophysical testing

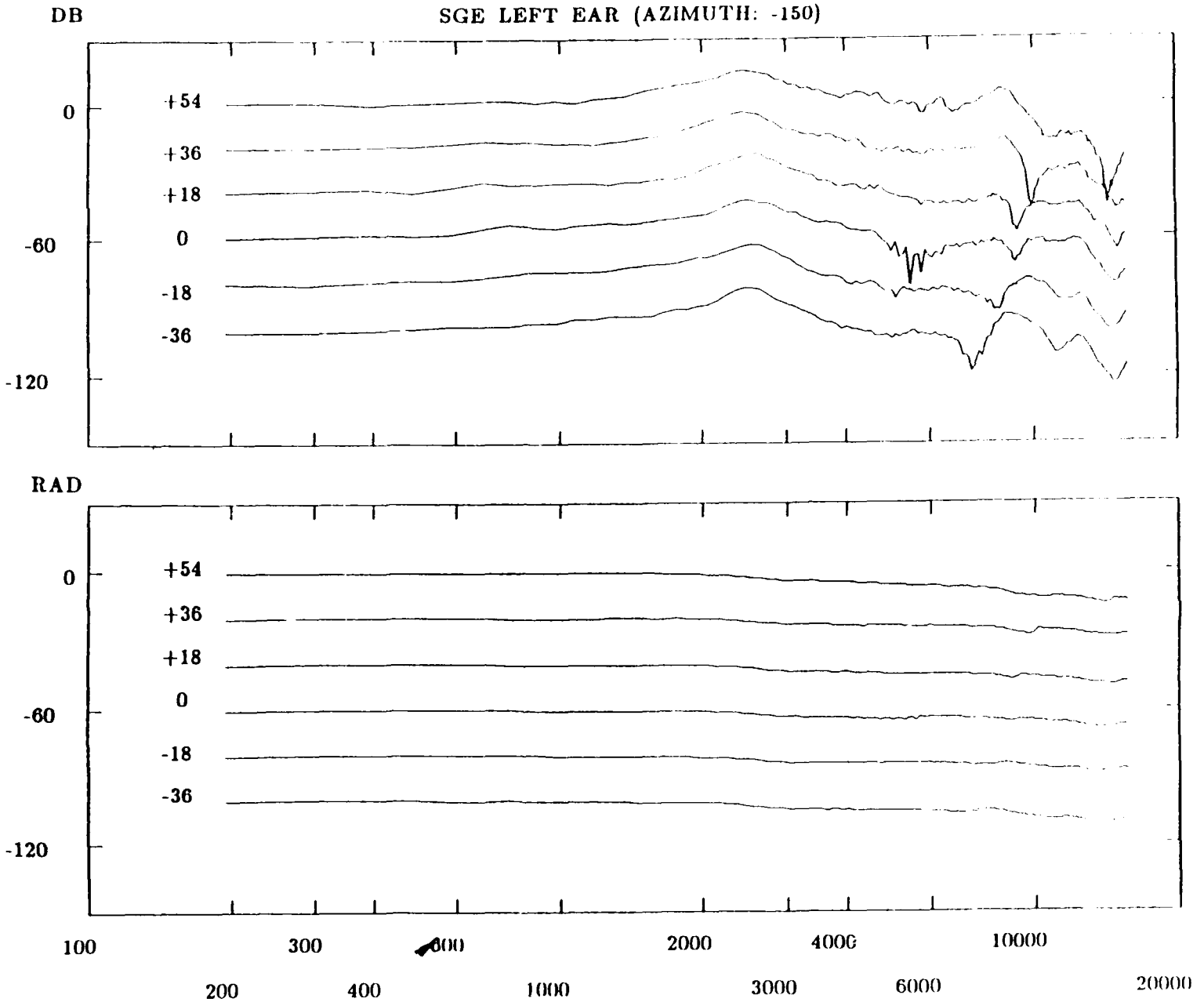
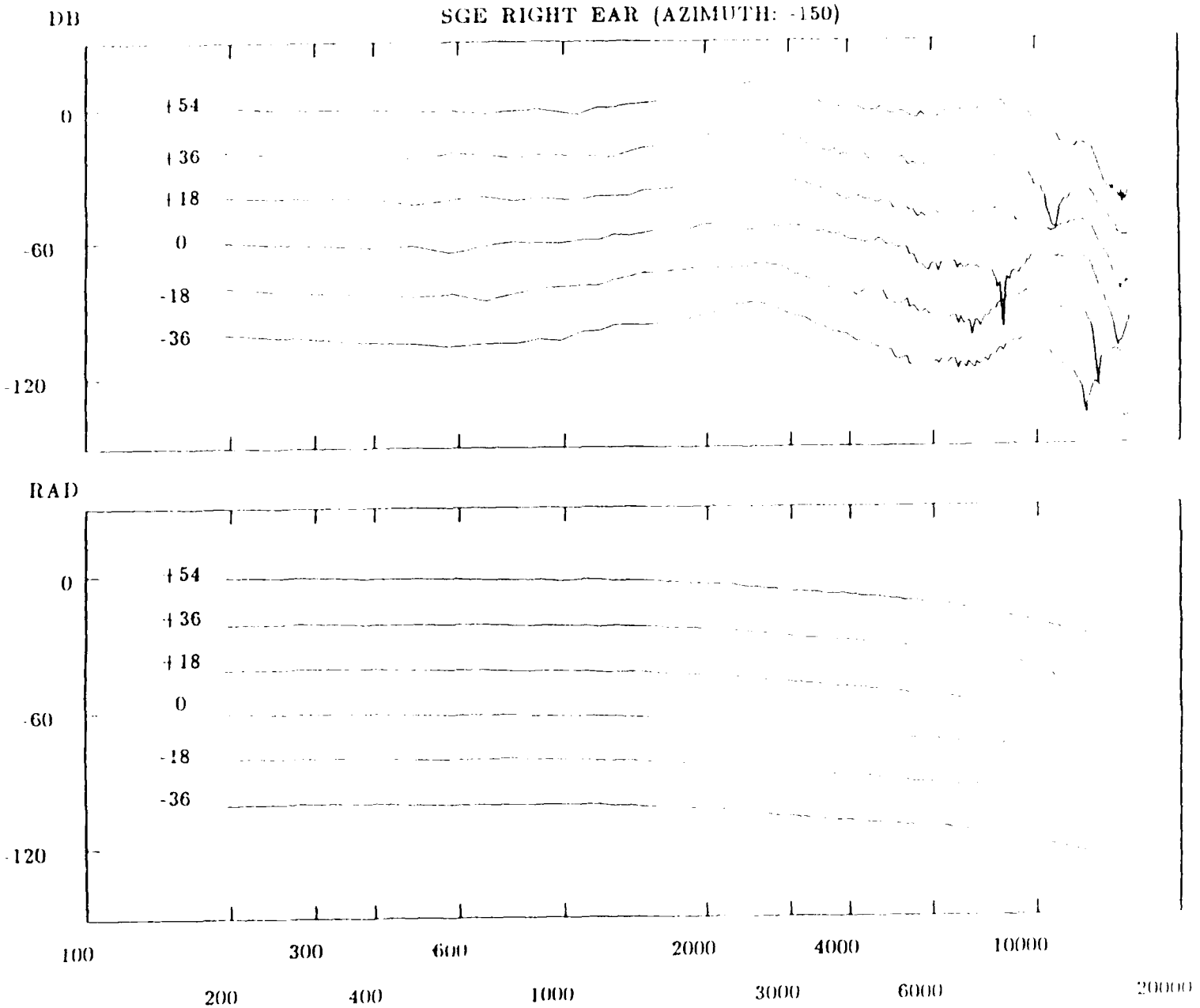


Figure 5. HRTF measurements from subject SGE for a source at -150 degrees azimuth i.e., 150 degrees on the subject's left side) and+54, +36, +18, 0, -18, -36 degrees elevation. The left ear HRTFis plotted in panel (a) and the right ear, in panel (b). The upper panel in (a) and (b) is the magnitude function in dB coordinates and the lower panel is the"unwrapped" phase function in radian coordinates The functions for the six elevations are displace dvertically by 20 dB for visibility. The measurements were obtained using a periodic pseudorandom noise


SGE RIGHT EAR (AZIMUTH: -150)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000

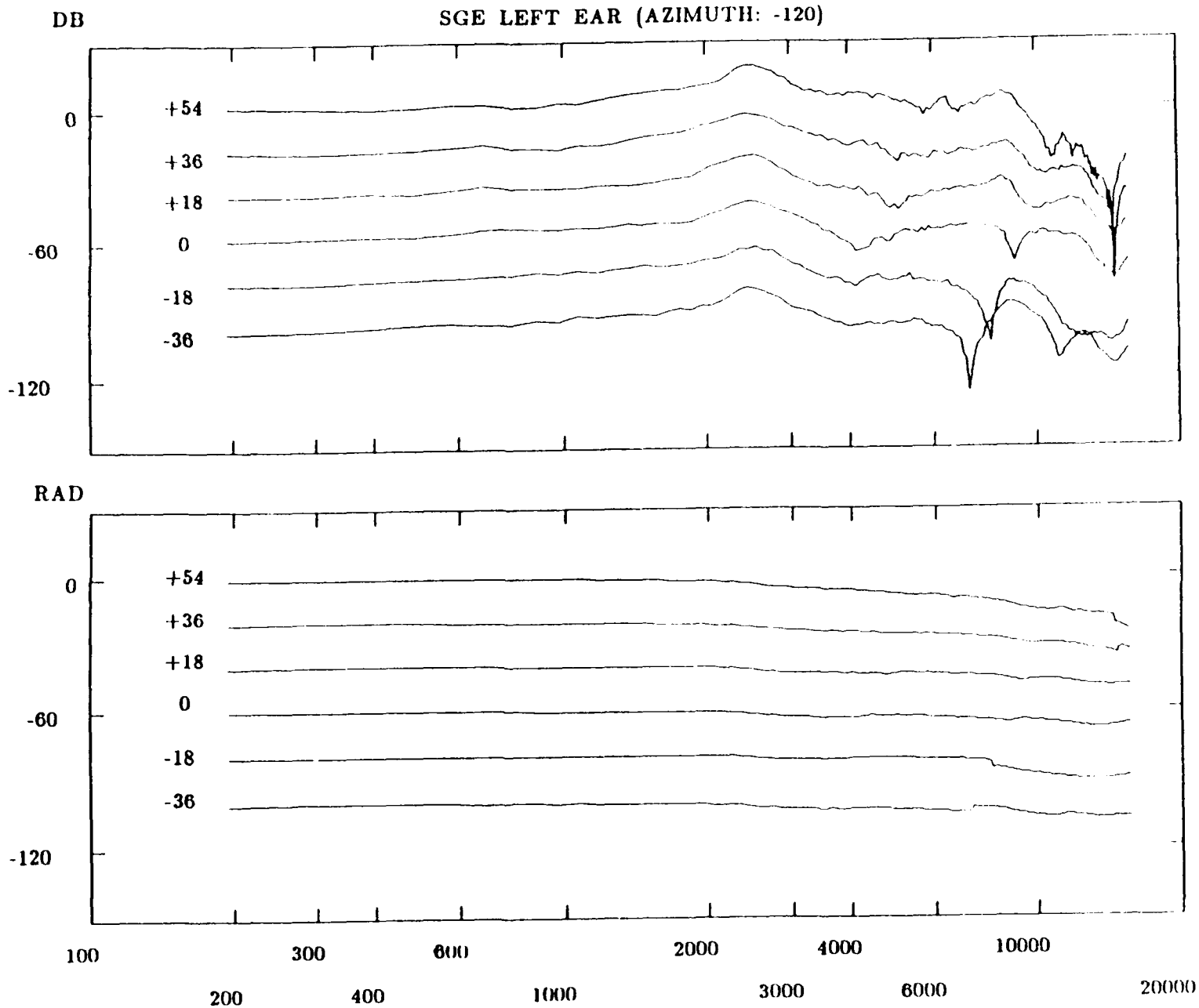
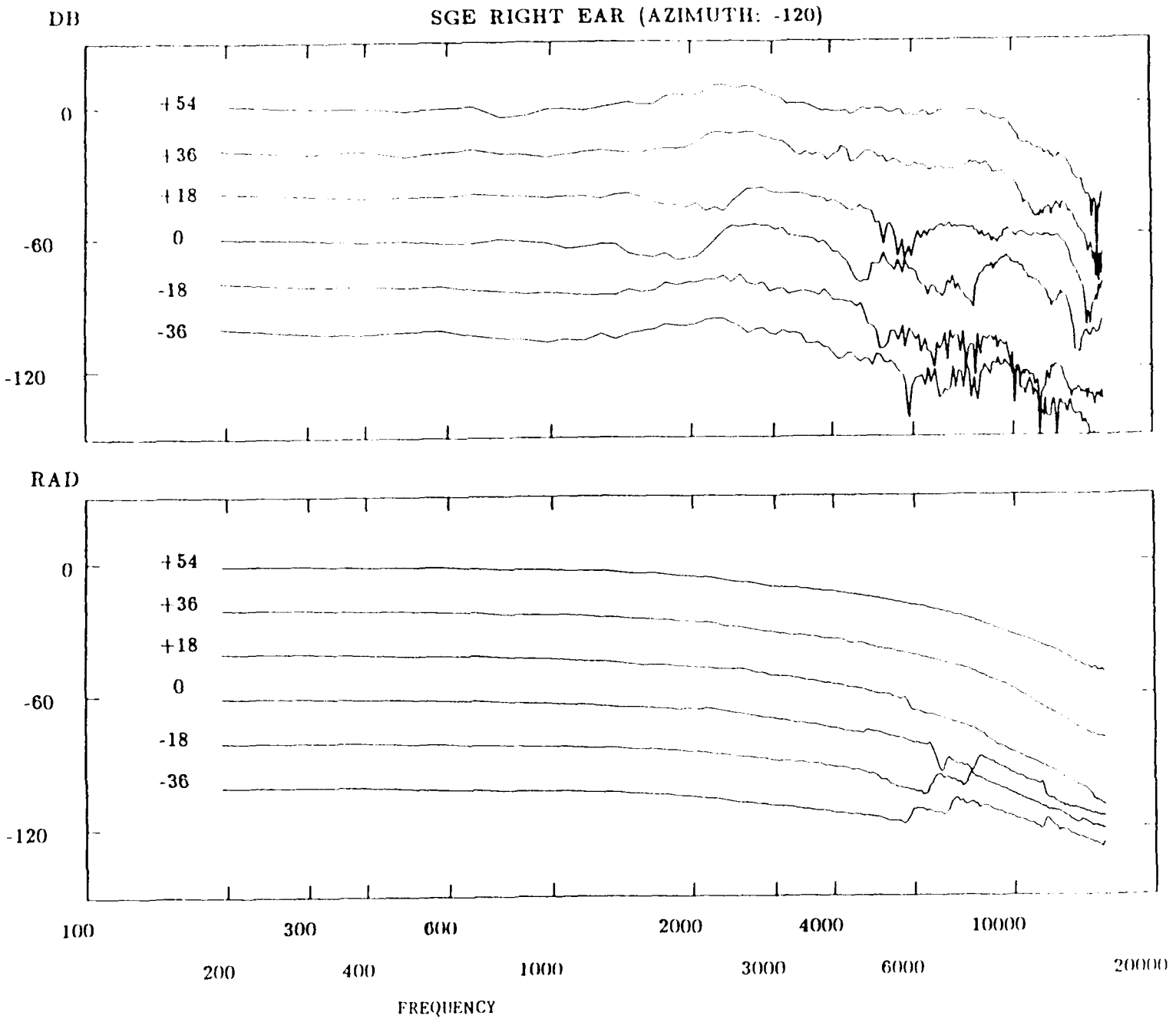


Figure 6. Same as Figure 5, except for a source at -120 degrees azimuth


SGE RIGHT EAR (AZIMUTH: -120)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

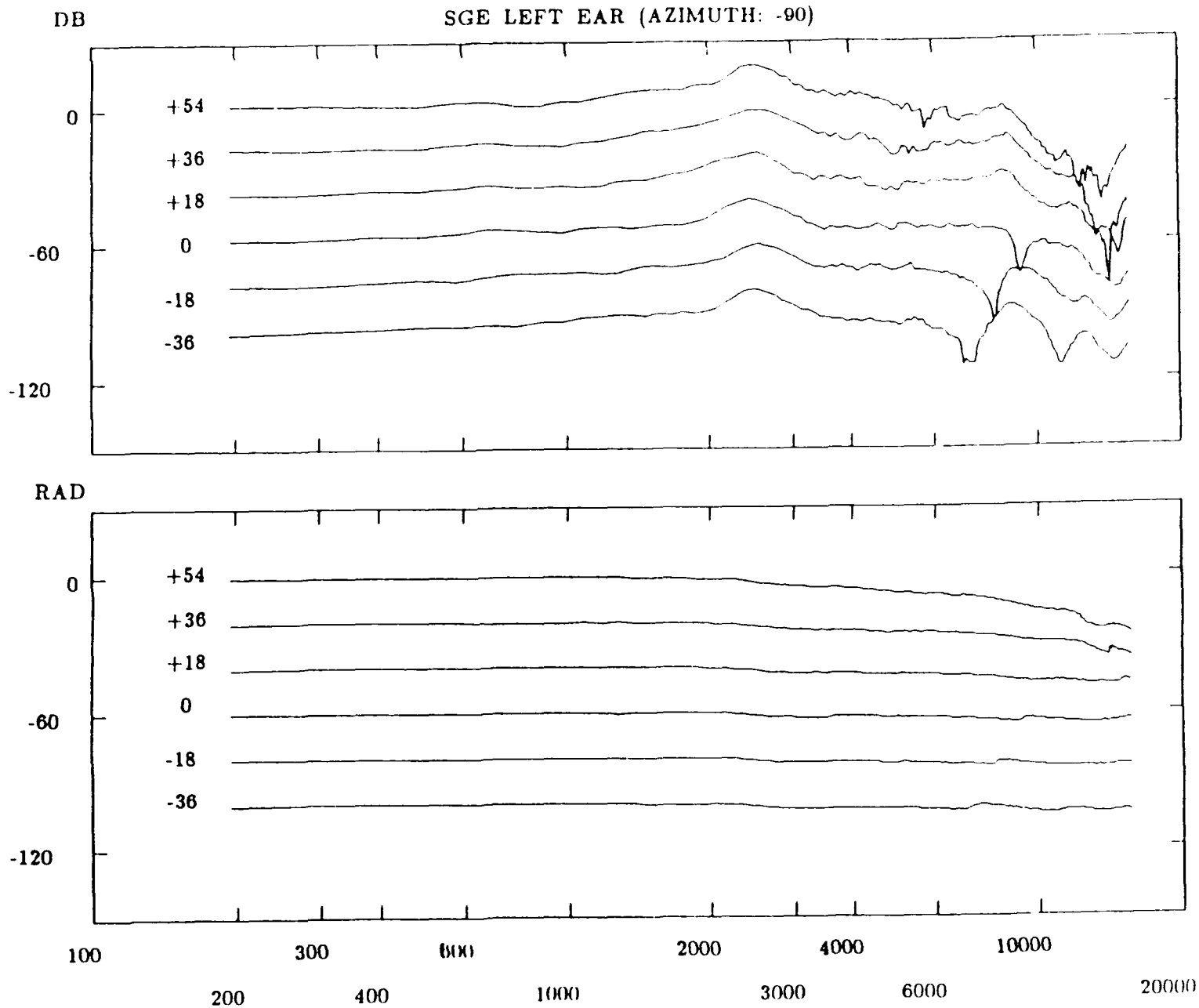
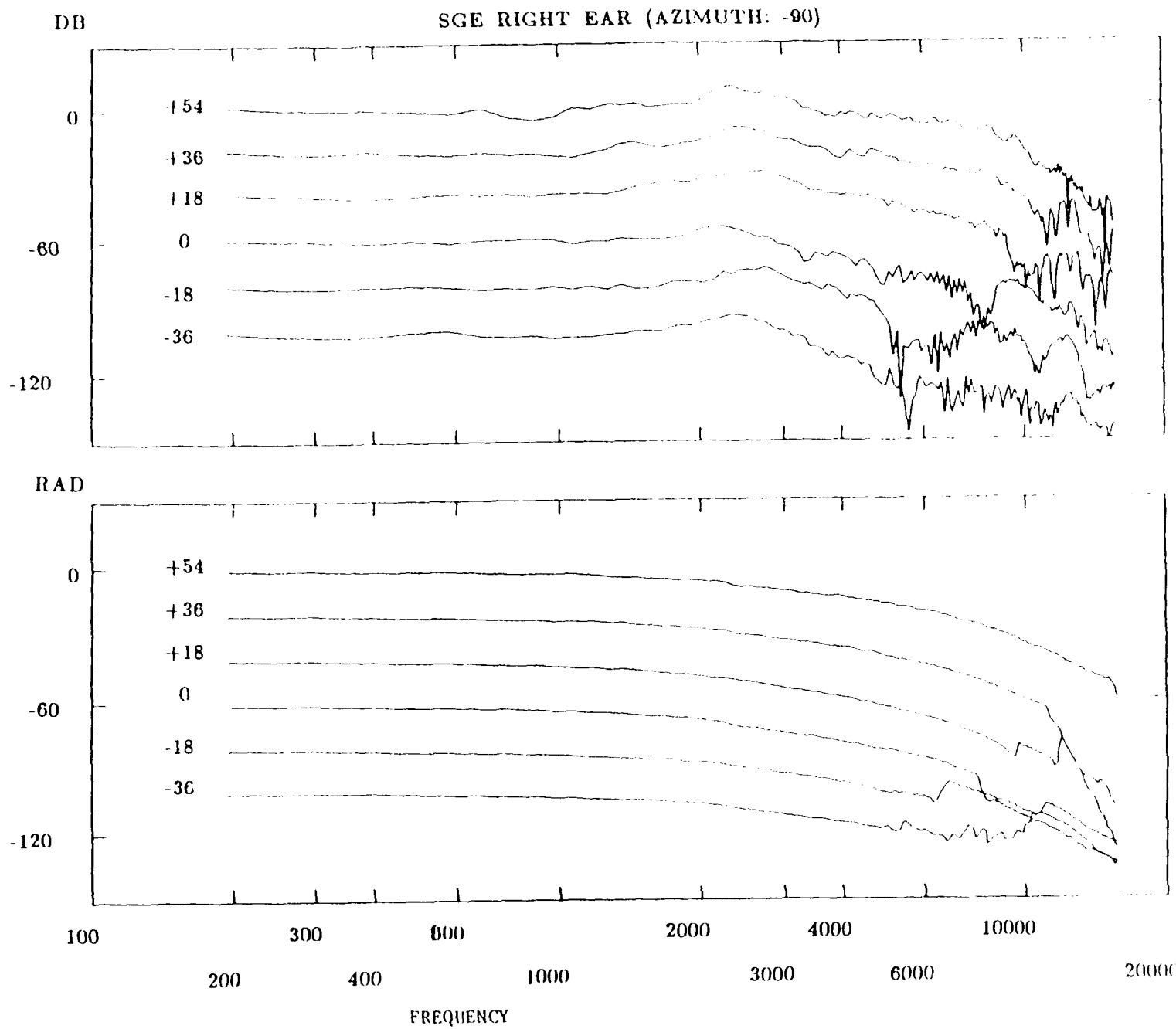


Figure 7. Same as Figure 5, except for a source at -90 degrees azimuth


SGE RIGHT EAR (AZIMUTH: -90)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

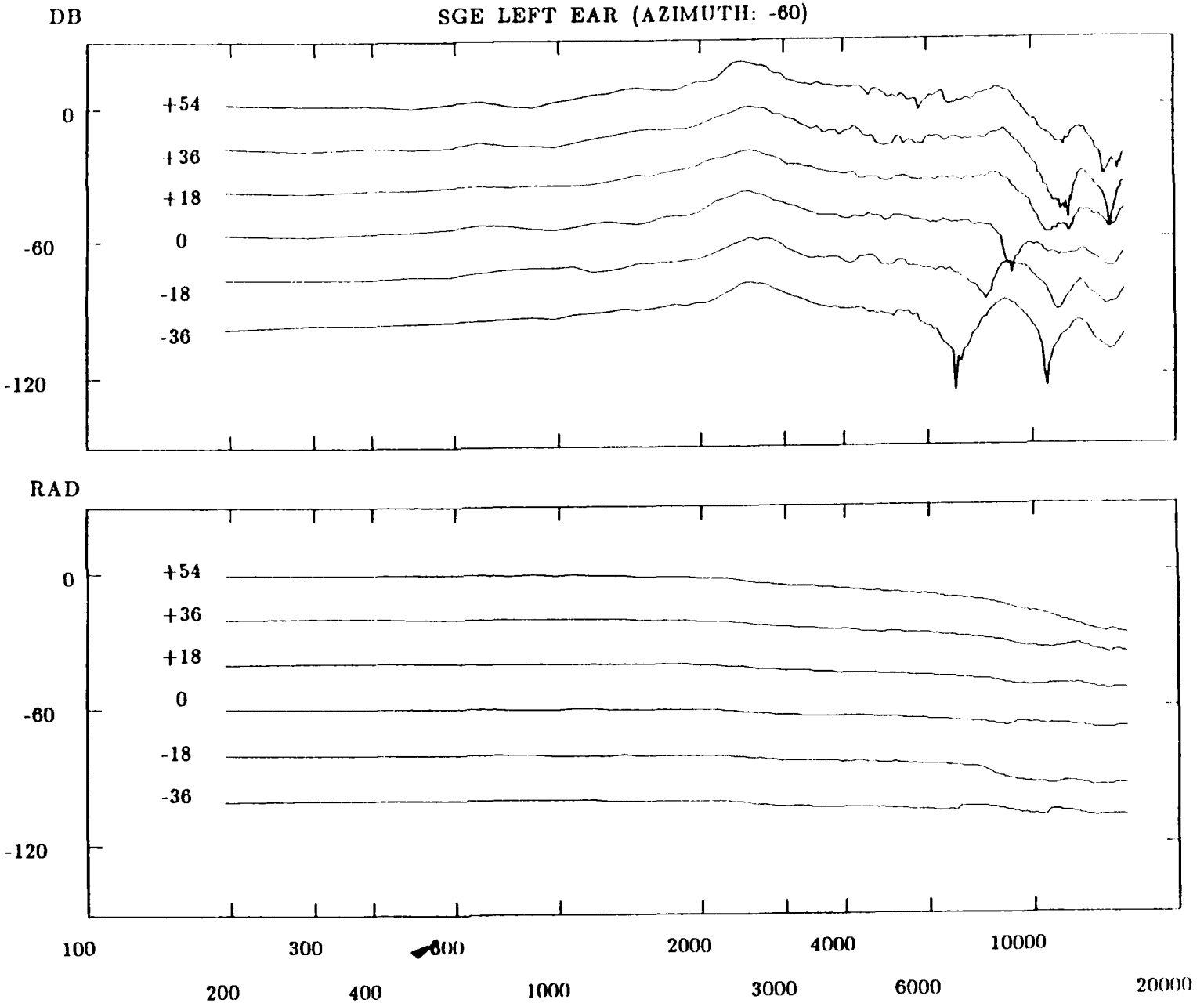
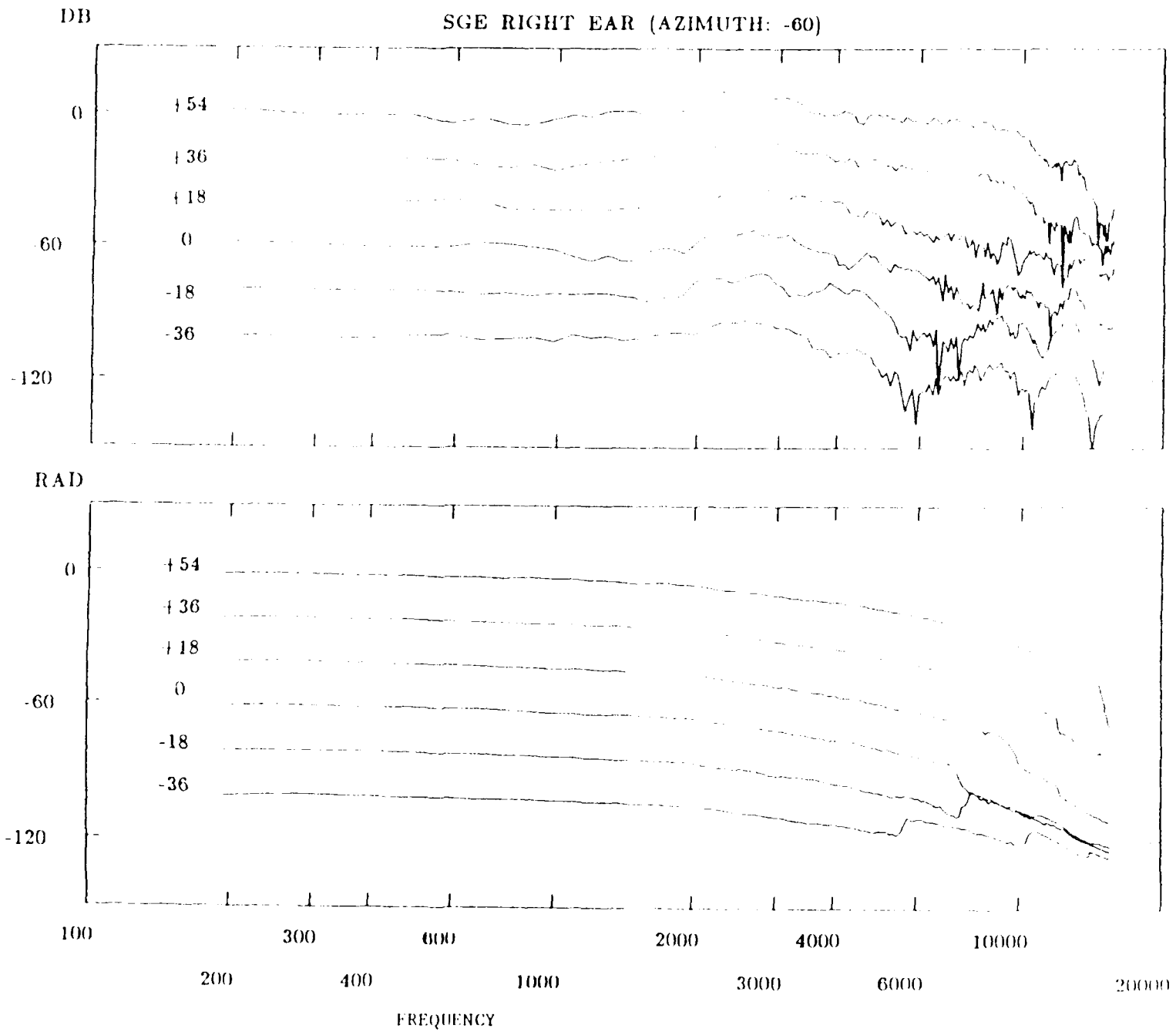


Figure 8. Same as Figure 5, except for a source at -60 degrees azimuth


SGE RIGHT EAR (AZIMUTH: -60)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

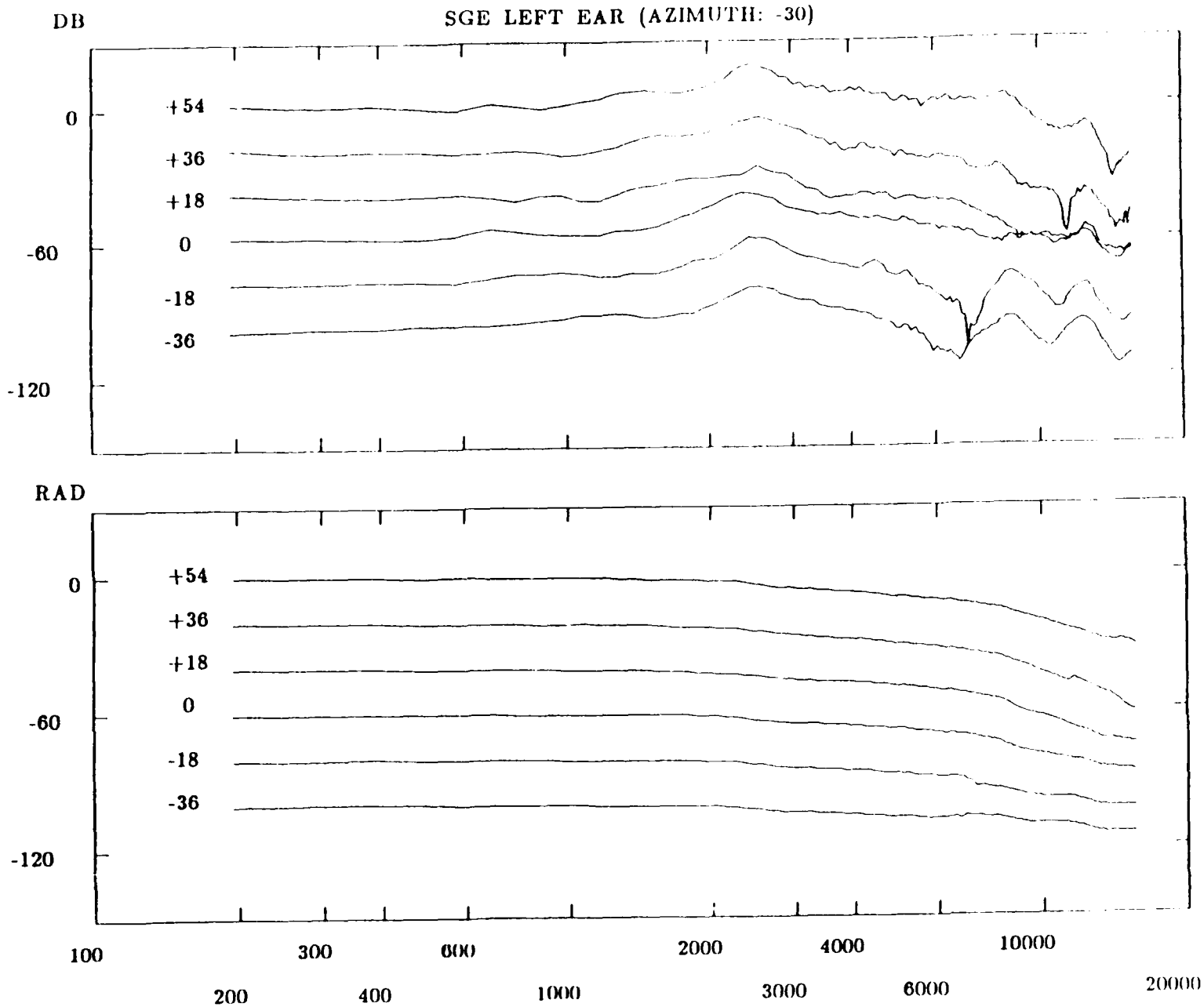
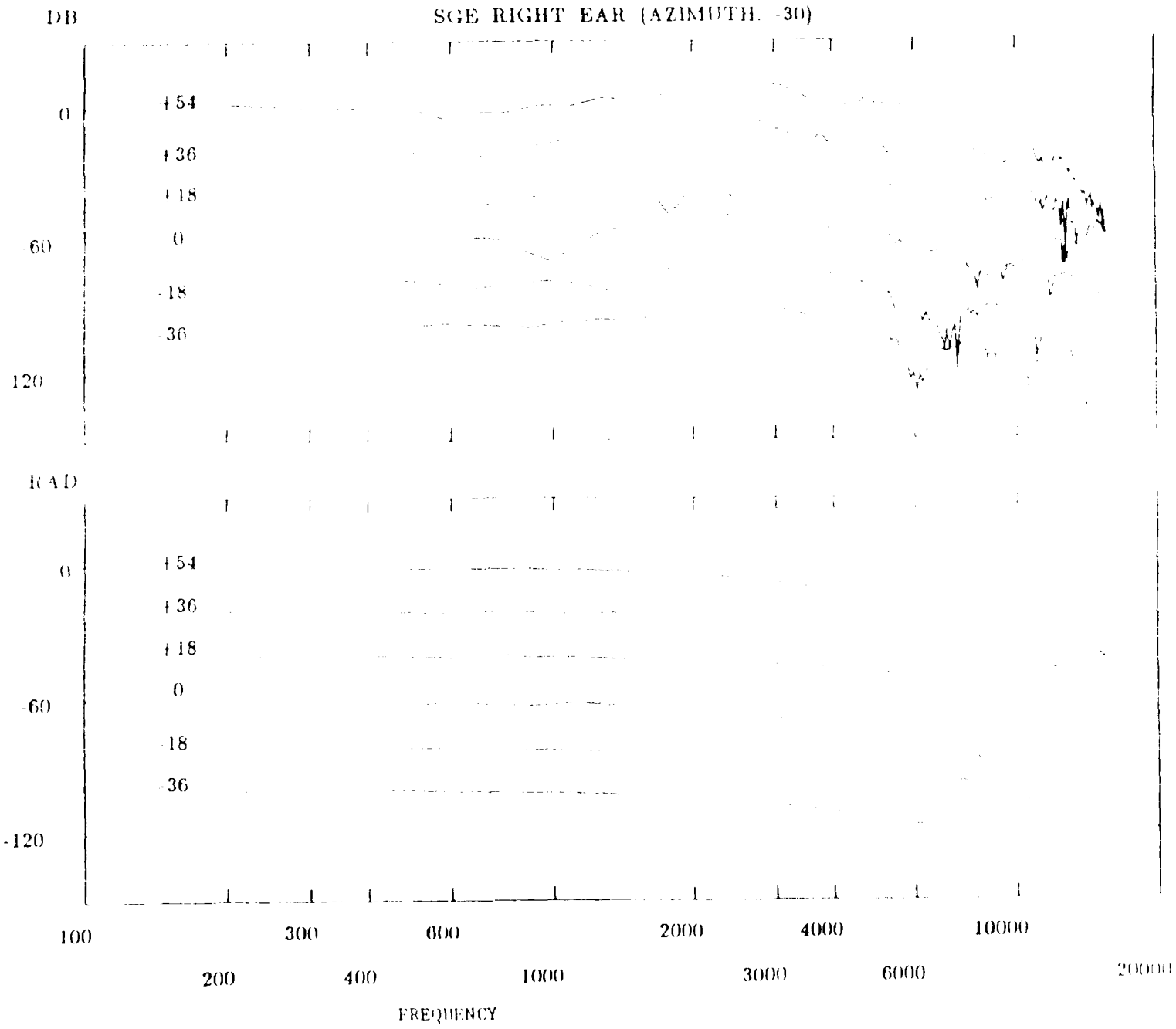


Figure 9. Same as Figure 5, except for a source at -30 degrees azimuth


DB
SGE RIGHT EAR (AZIMUTH -30)
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

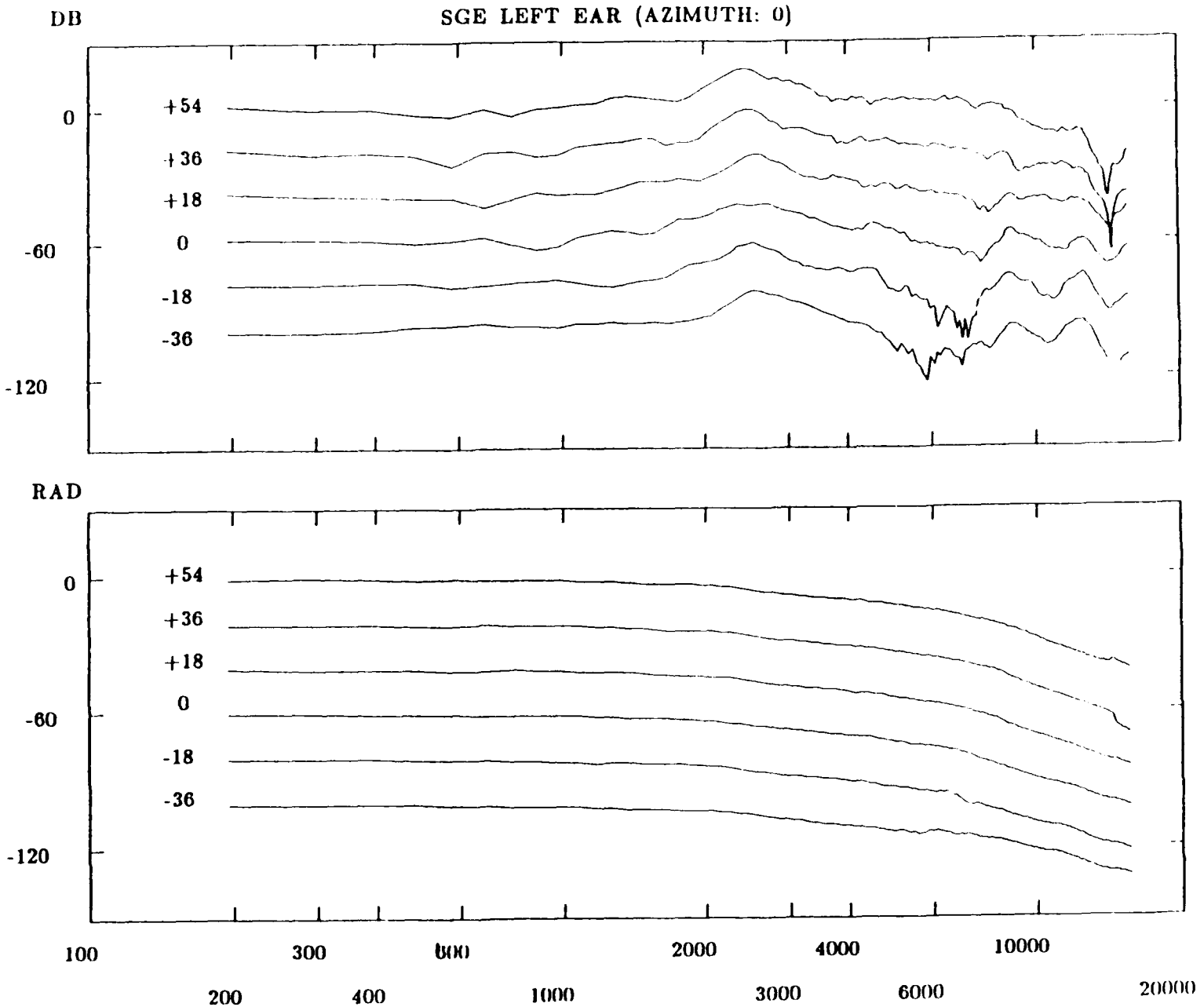
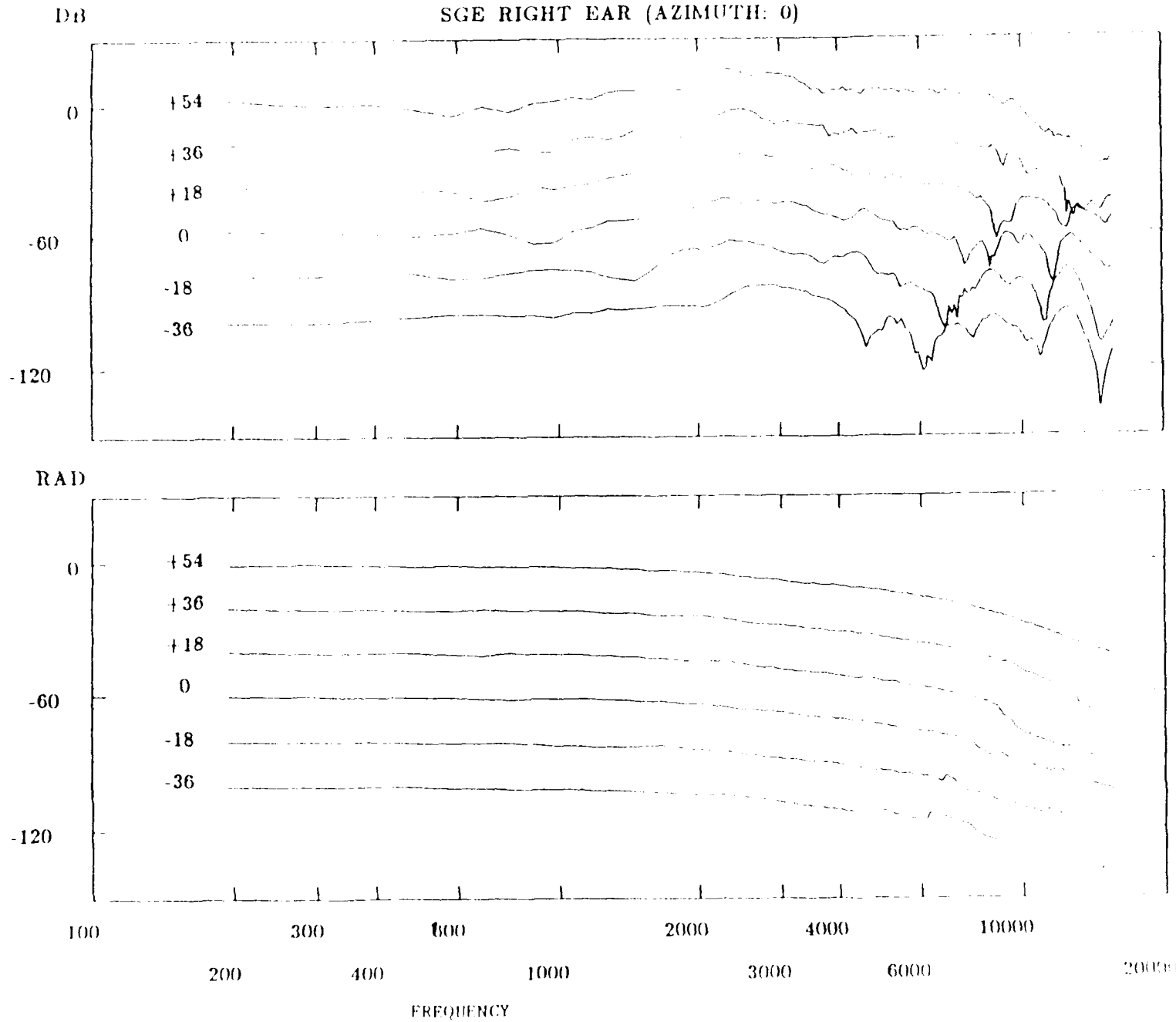


Figure 10. Same as Figure 5, except for a source at 0 degrees azimuth


DB
SGE RIGHT EAR (AZIMUTH: 0)
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
FREQUENCY

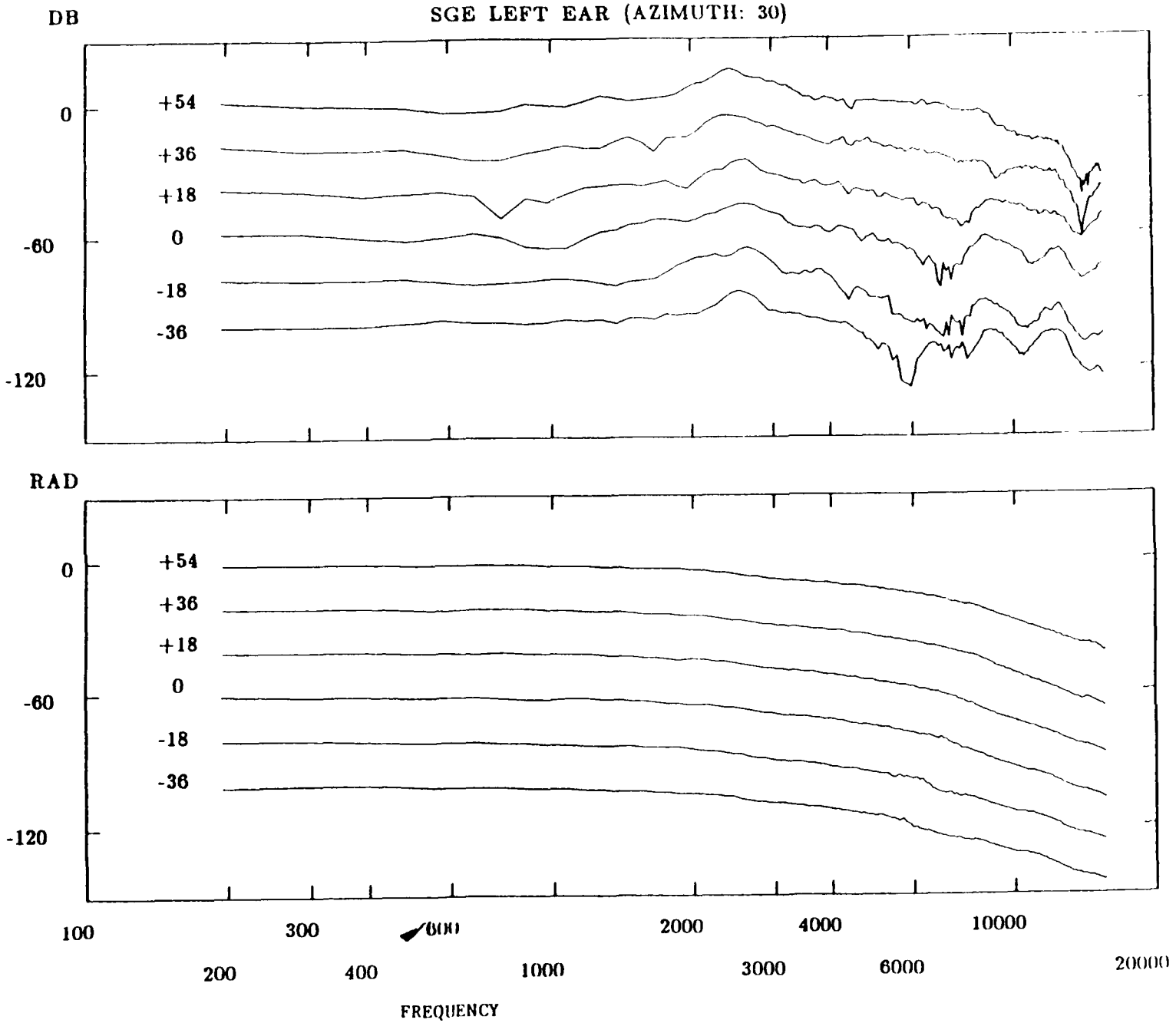
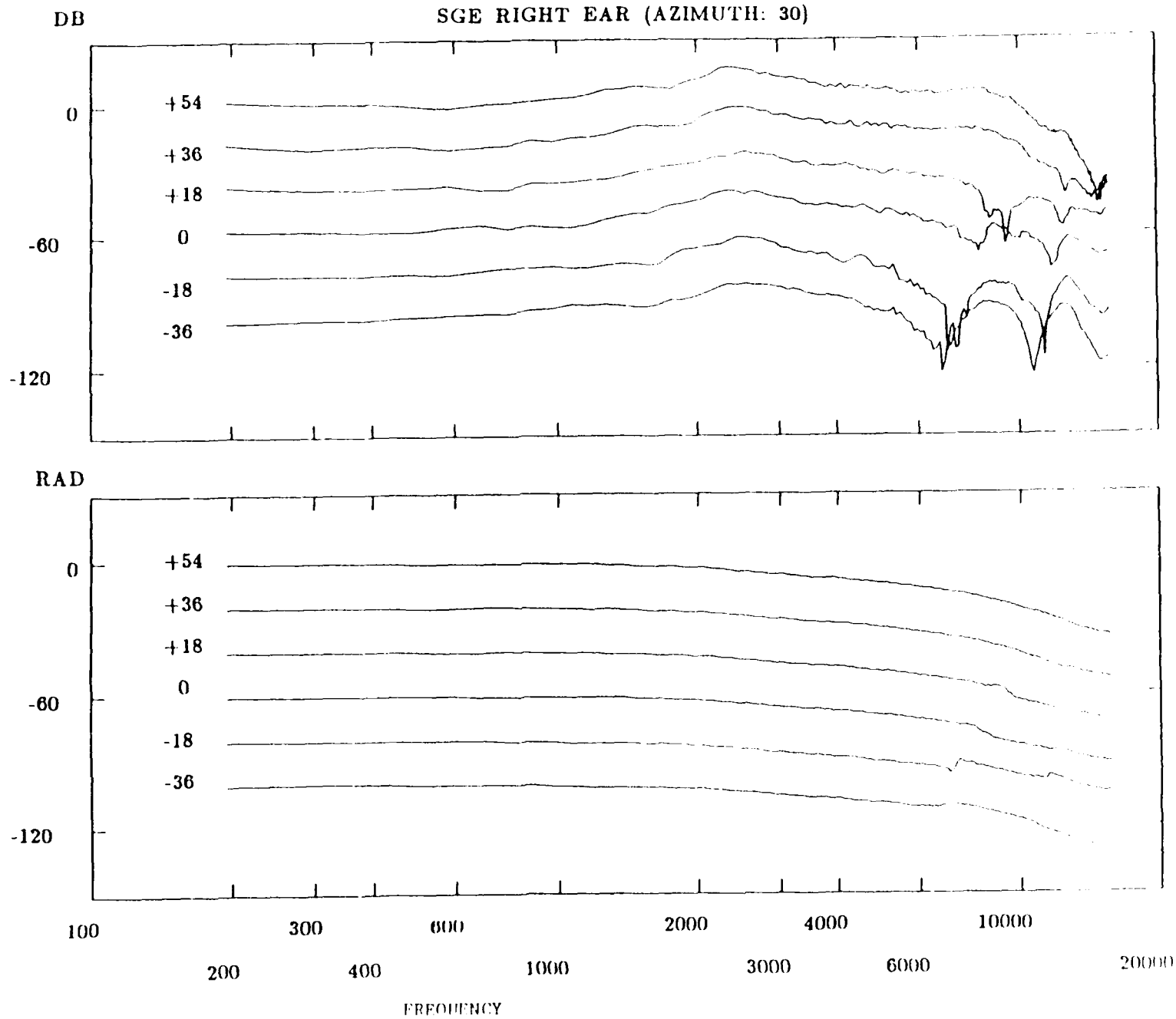


Figure 11. Same as Figure 5, except for a source at 30 degrees azimuth


SGE RIGHT EAR (AZIMUTH: 30)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

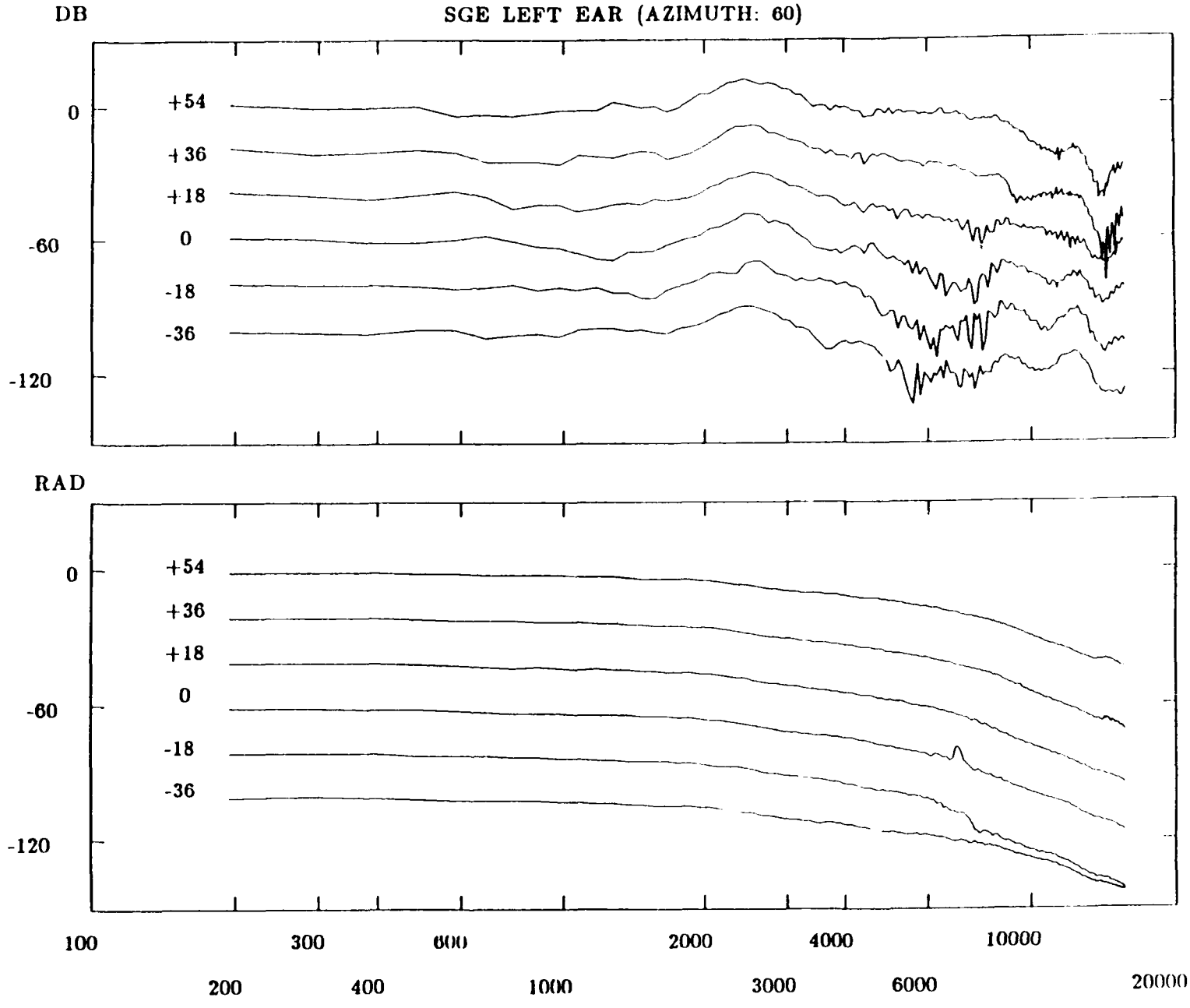
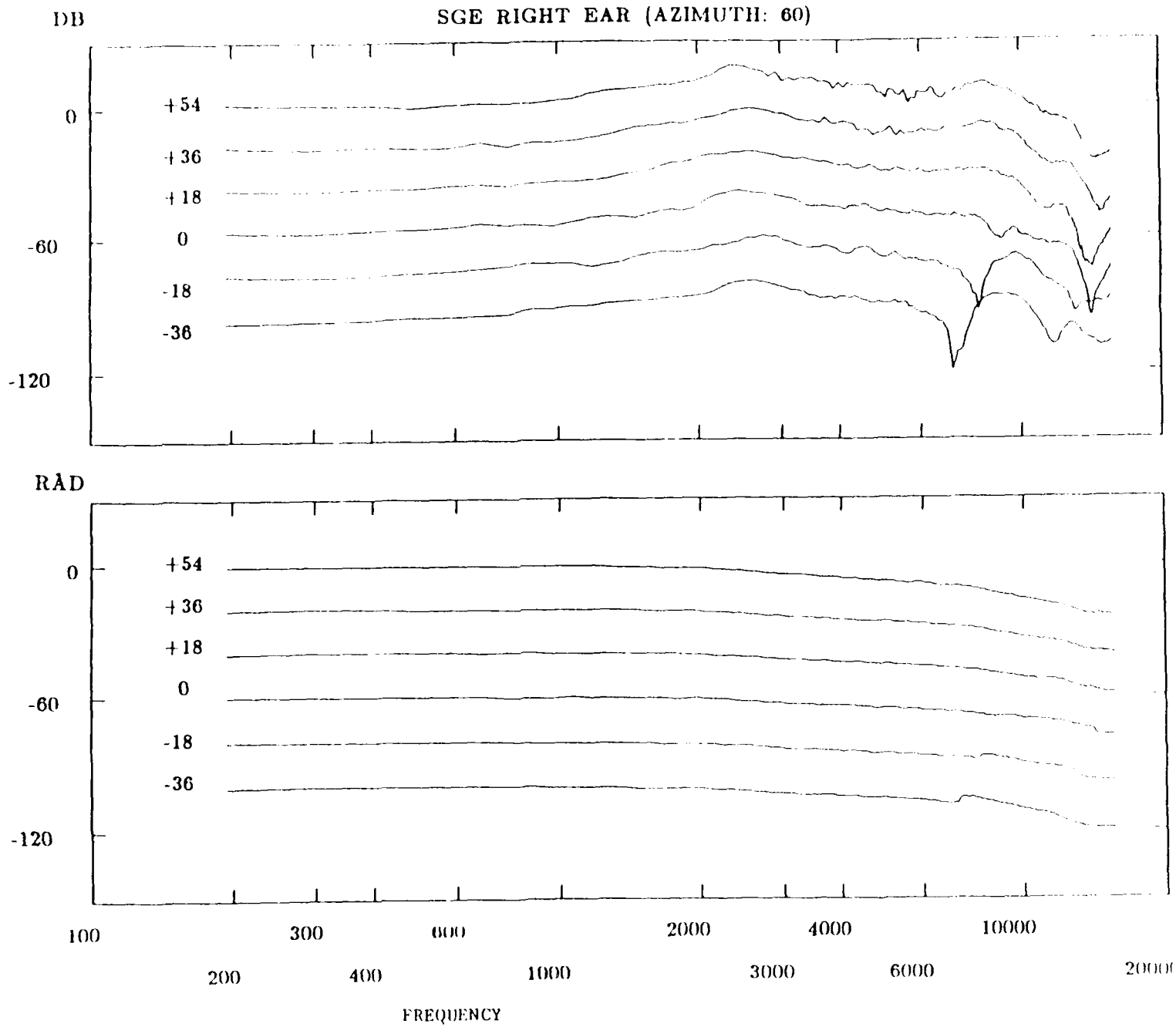


Figure 12. Same as Figure 5, except for a source at 60 degrees azimuth


SGE RIGHT EAR (AZIMUTH: 60)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

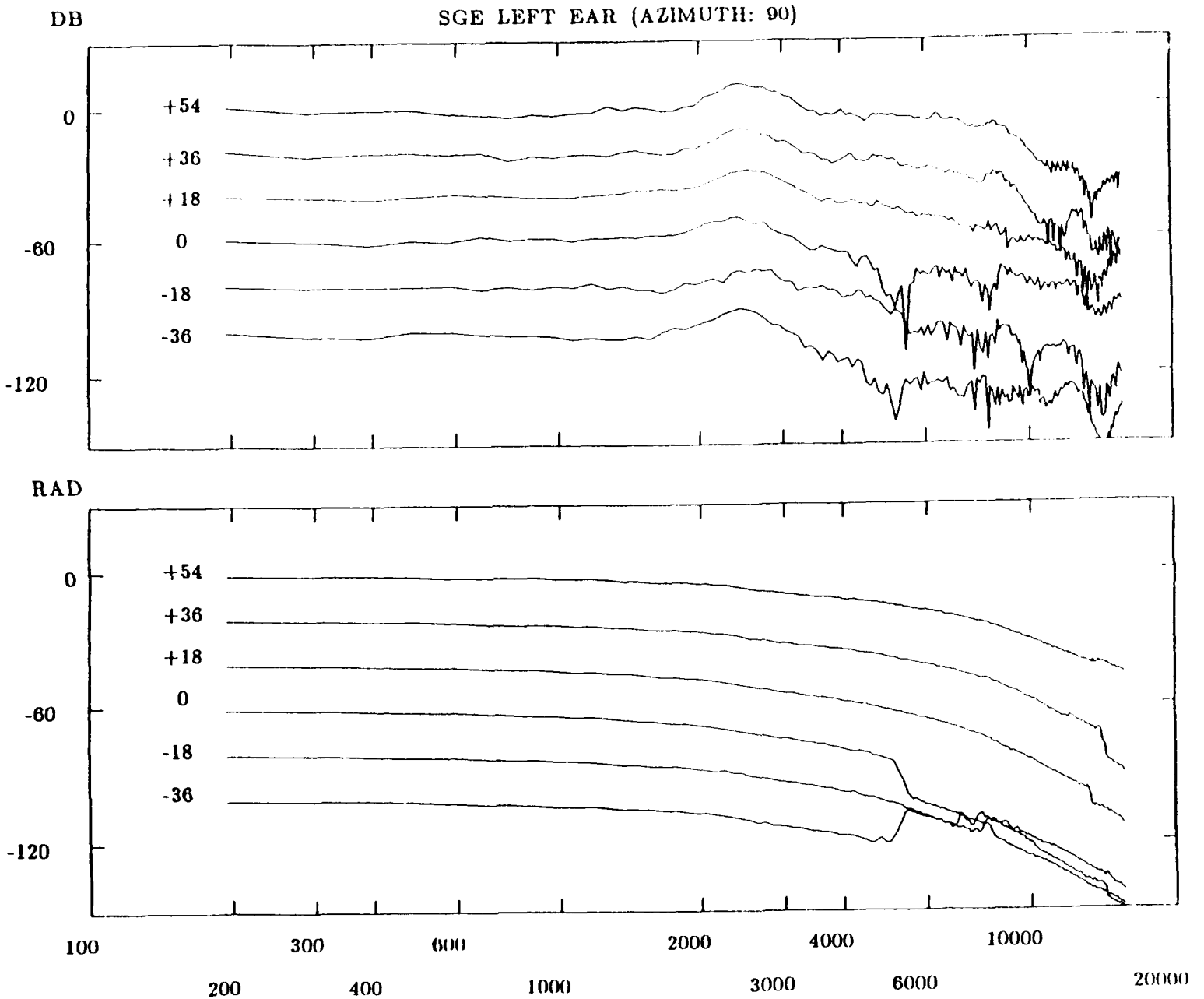
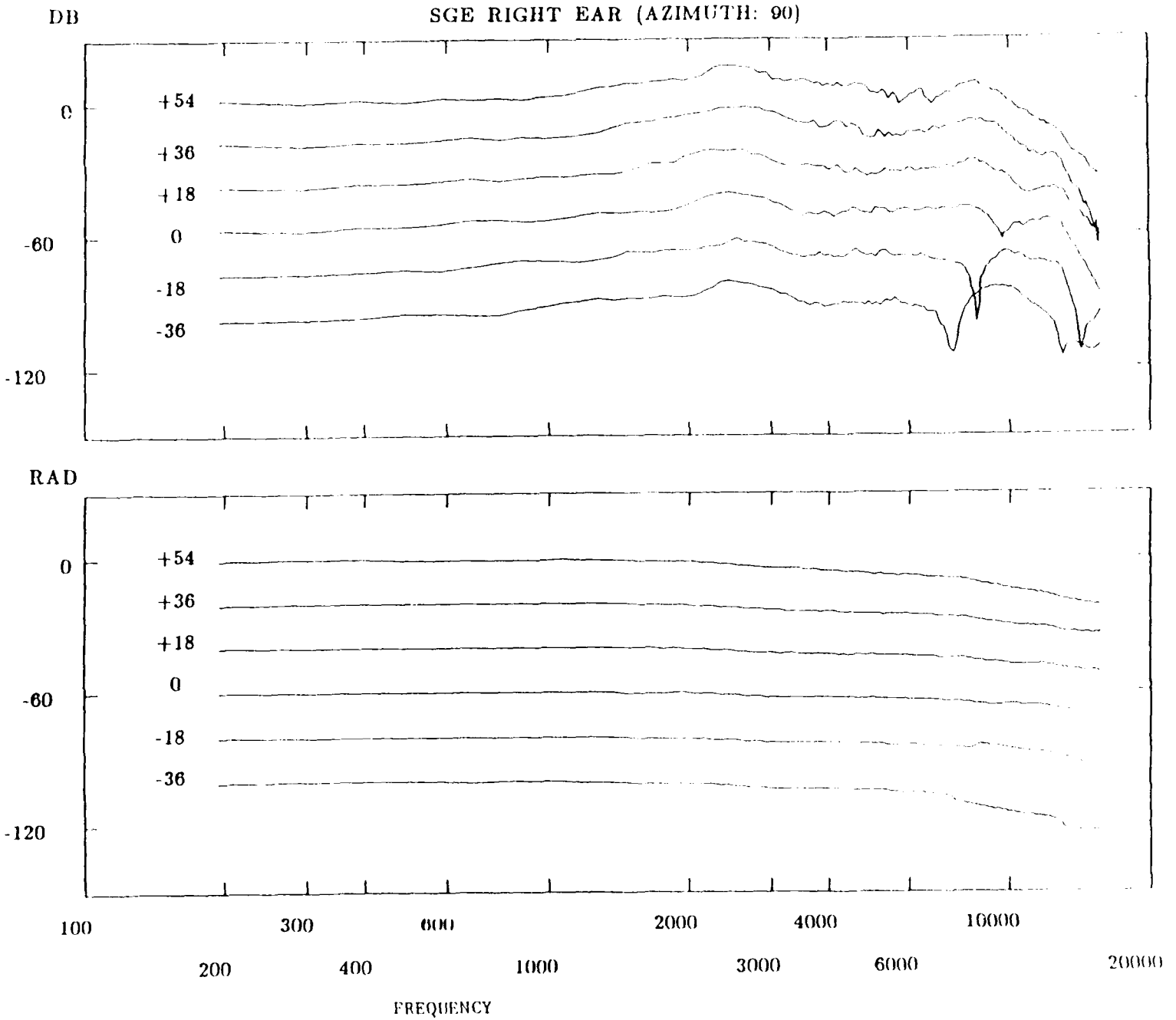


Figure 13. Same as Figure 5, except for a source at 90 degrees azimuth


SGE RIGHT EAR (AZIMUTH: 90)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

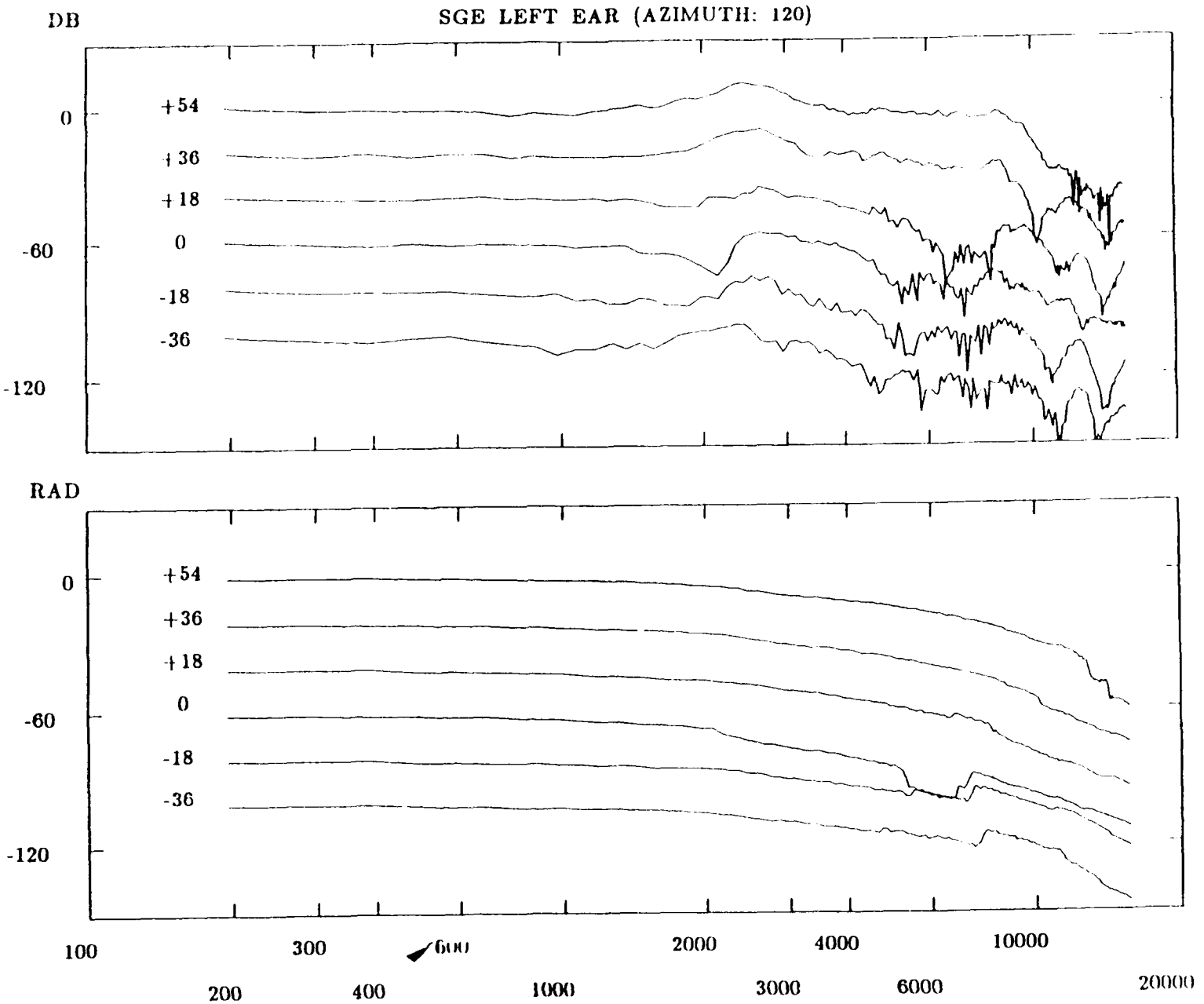
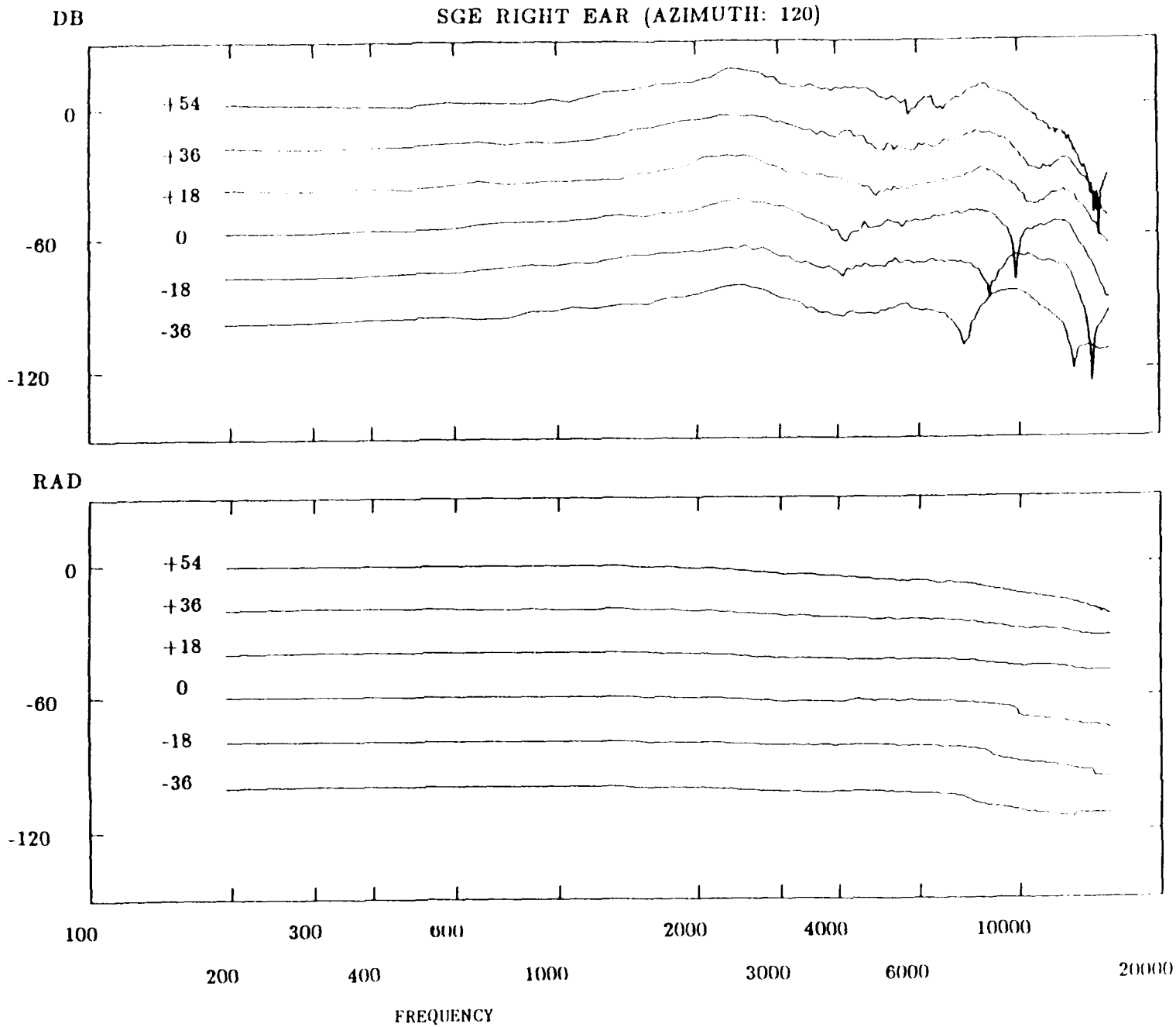


Figure 14. Same as Figure 5, except for a source at 120 degrees azimuth


SGE RIGHT EAR (AZIMUTH: 120)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

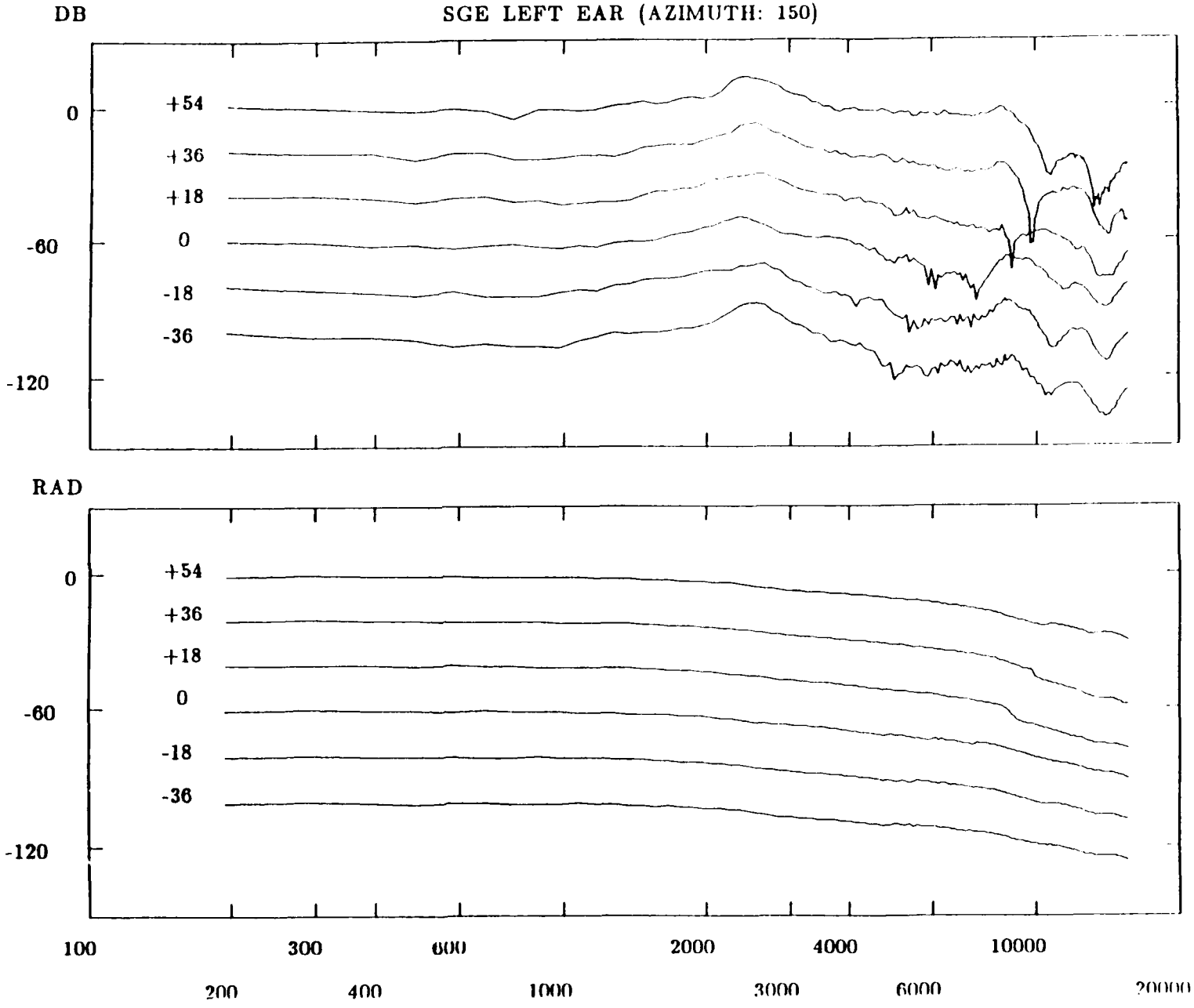
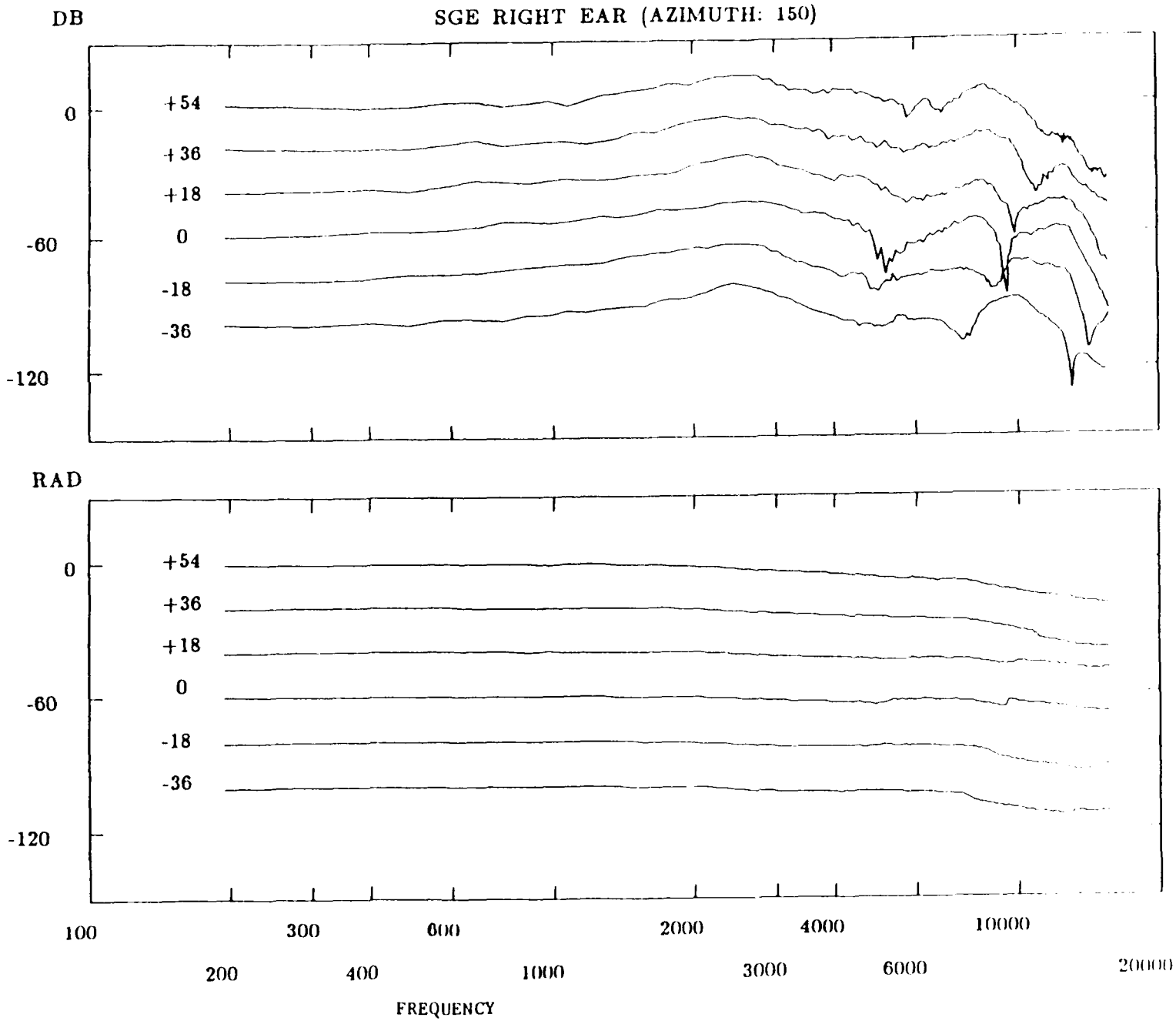


Figure 15. Same as Figure 5, except for a source at 150 degrees azimuth


SGE RIGHT EAR (AZIMUTH: 150)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

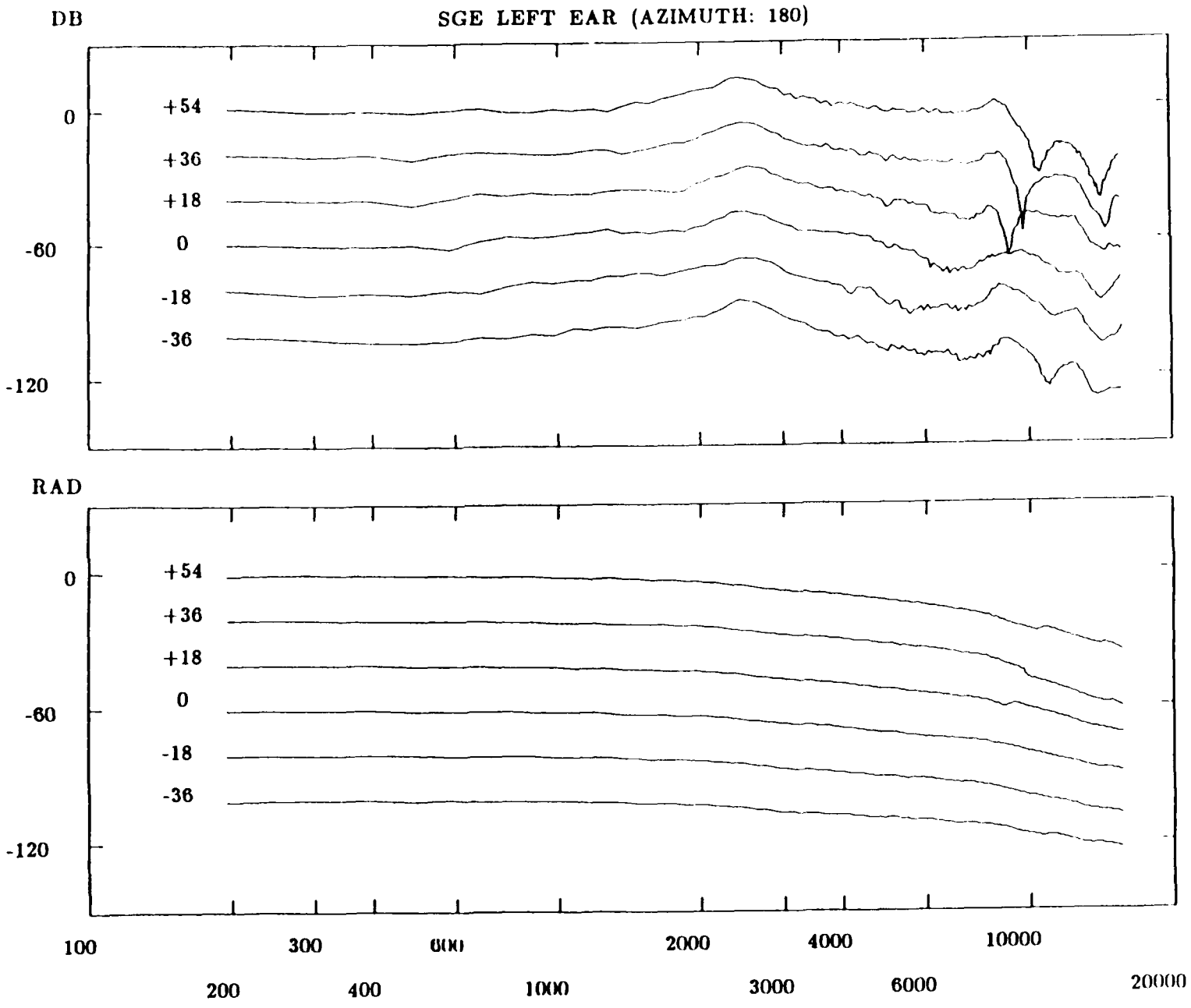
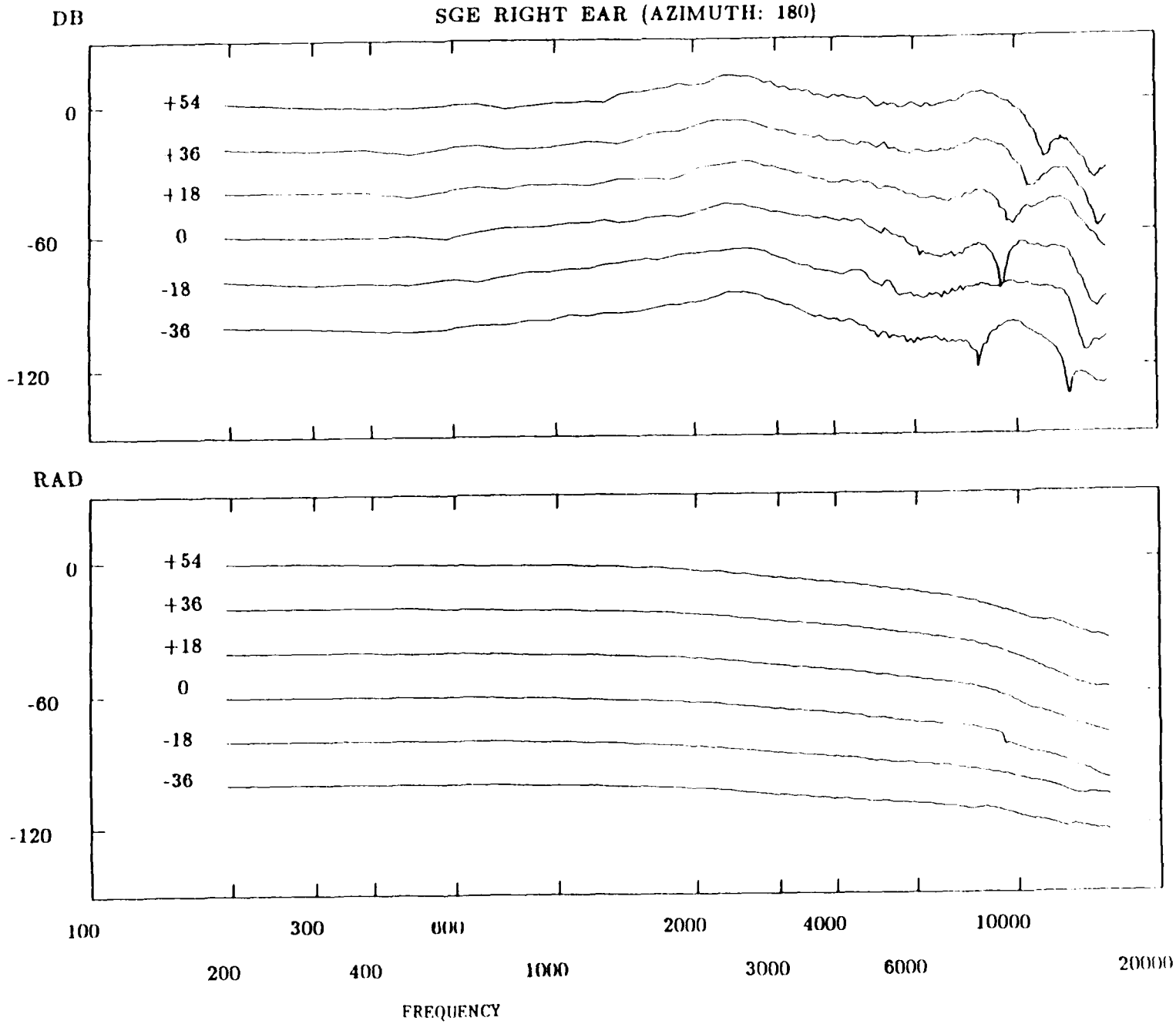


Figure 16. Same as Figure 5, except for a source at 180 degrees azimuth


SGE RIGHT EAR (AZIMUTH: 180)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

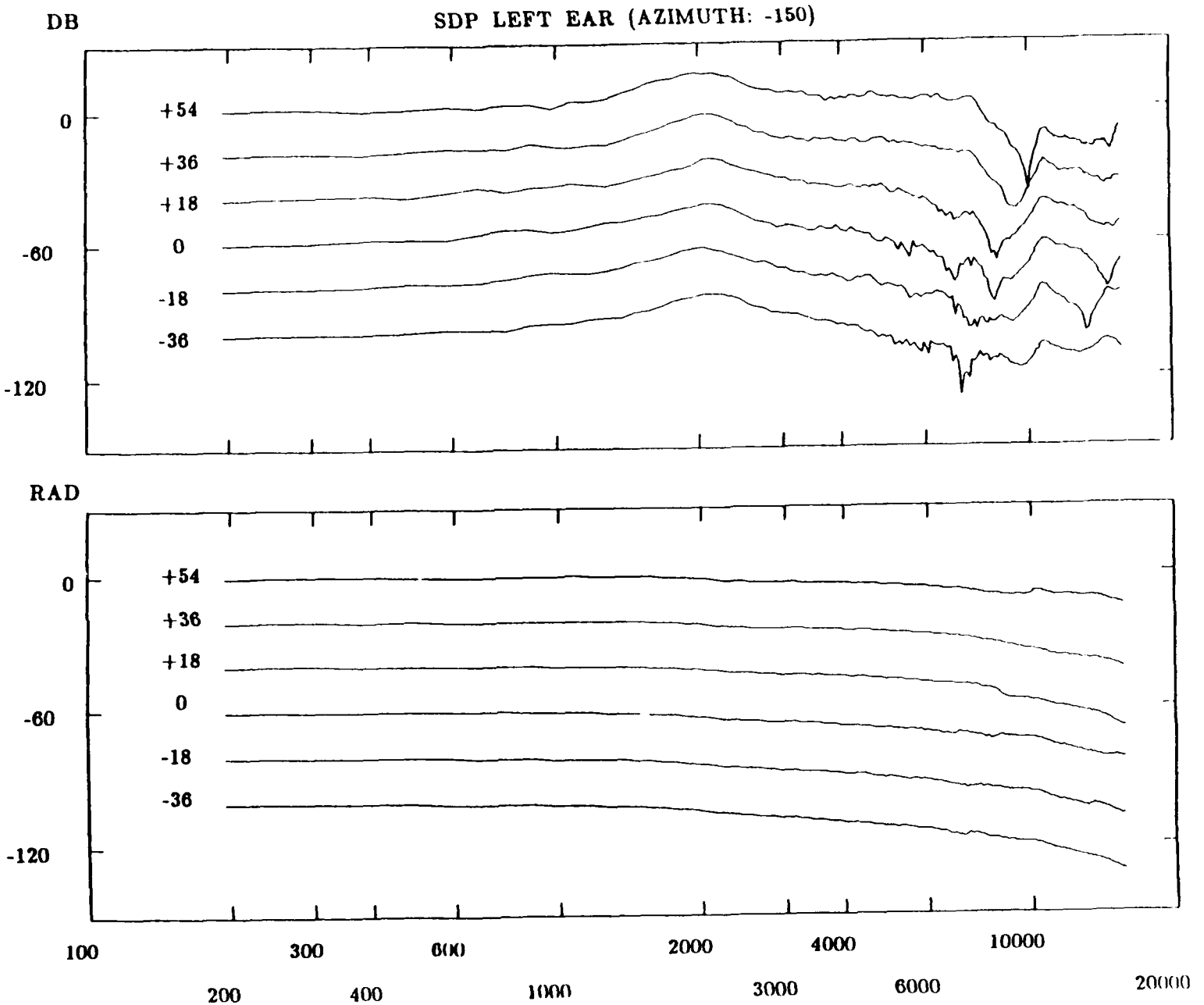
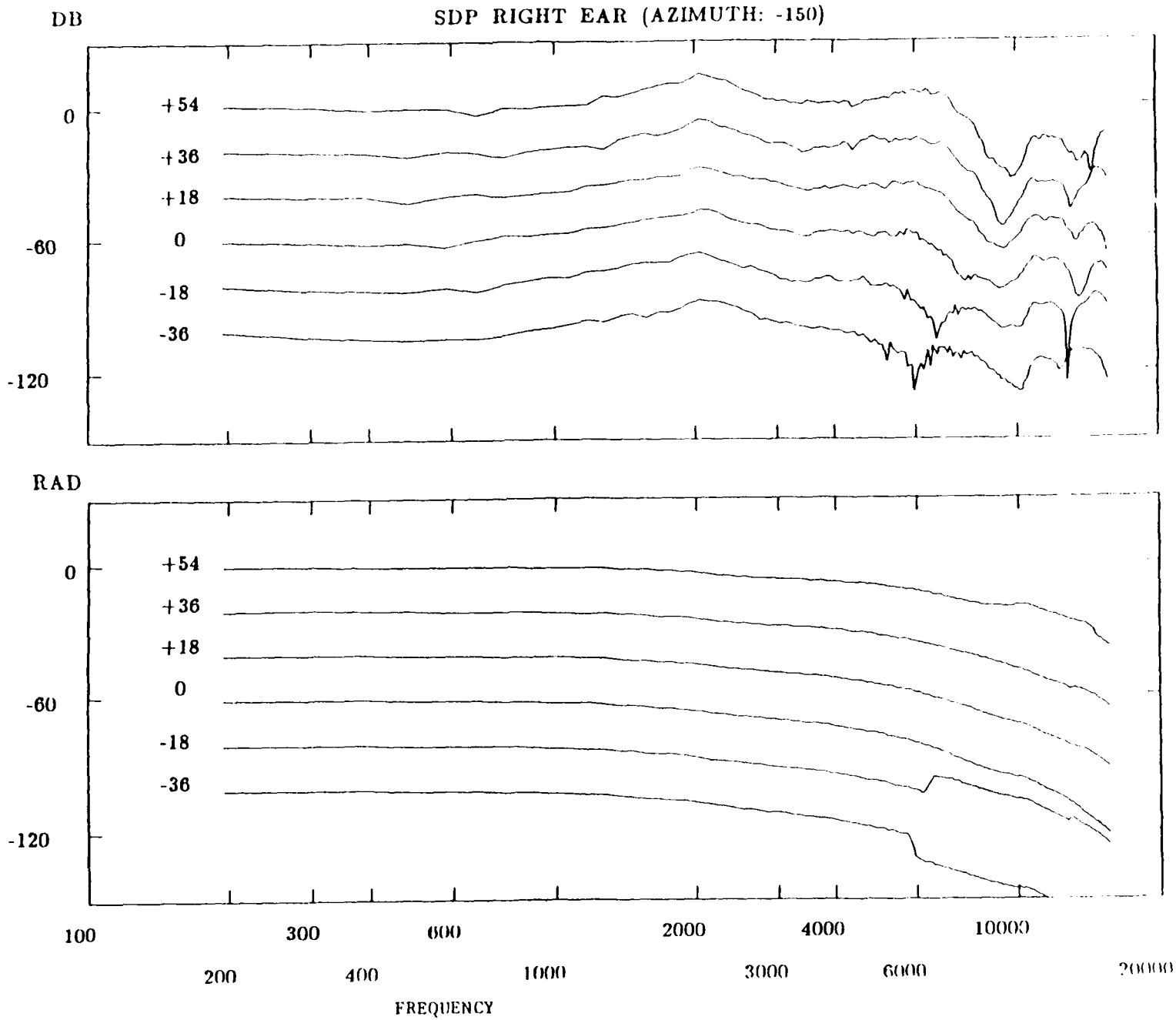


Figure 17. HRTF measurements from subject SDP for a source at -150 degrees azimuth (i.e., 150 degrees on the subject's left side) and +54, +36, +18, 0, -18, -36 degrees elevation. The left ear HRTF is plotted in panel (a) and the right ear, in panel (b). The upper panel in (a) and (b) is the magnitude function in dB coordinates and the lower panel is the "unwrapped" phase function in radian coordinates. The functions for the six elevations are displaced vertically by 20 dB for visibility. The measurements were obtained using a periodic pseudorandom noise


SDP RIGHT EAR (AZIMUTH: -150)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

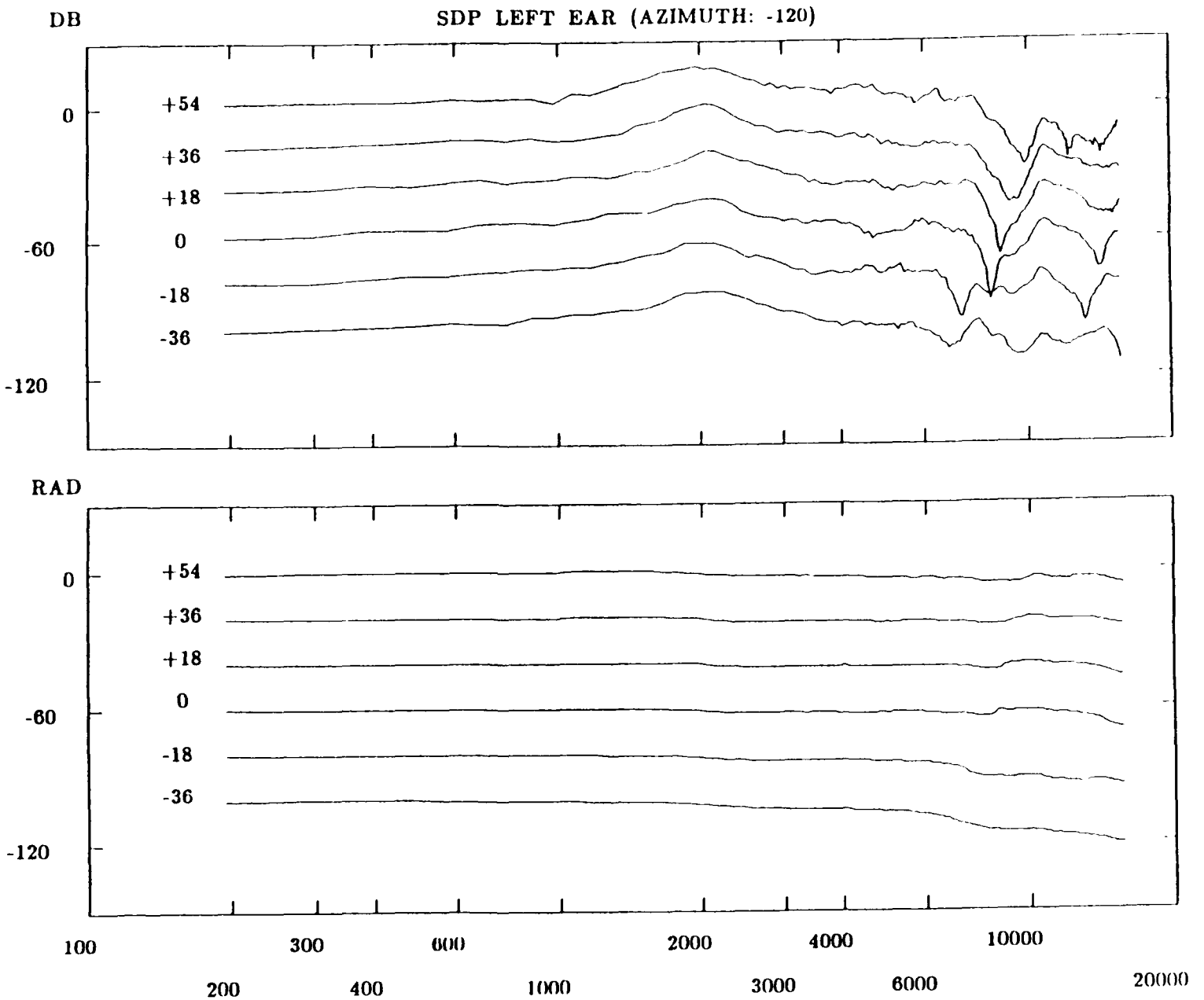
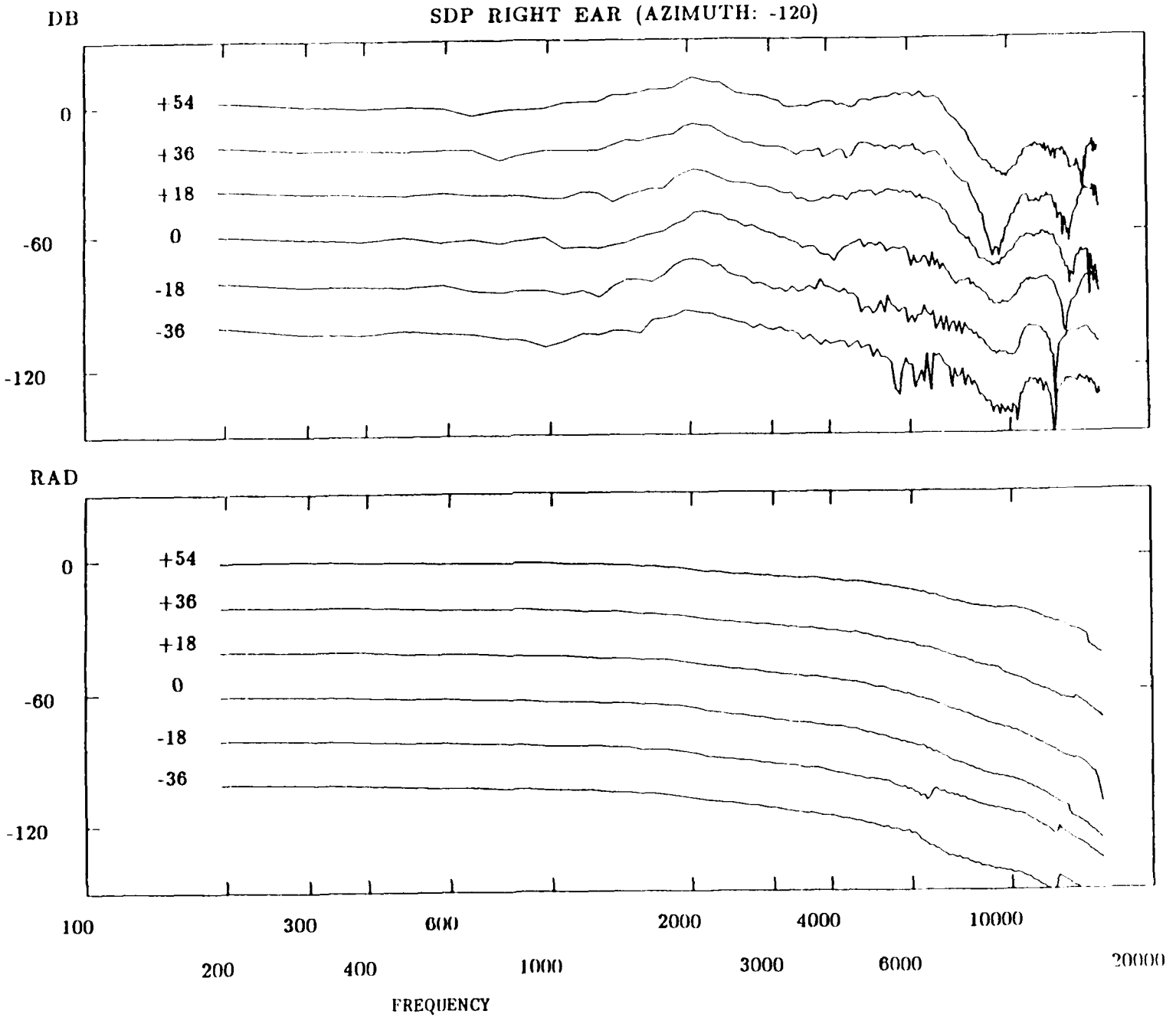


Figure 18. Same as Figure 17, except for a source at -120 degrees azimuth


SDP RIGHT EAR (AZIMUTH: -120)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

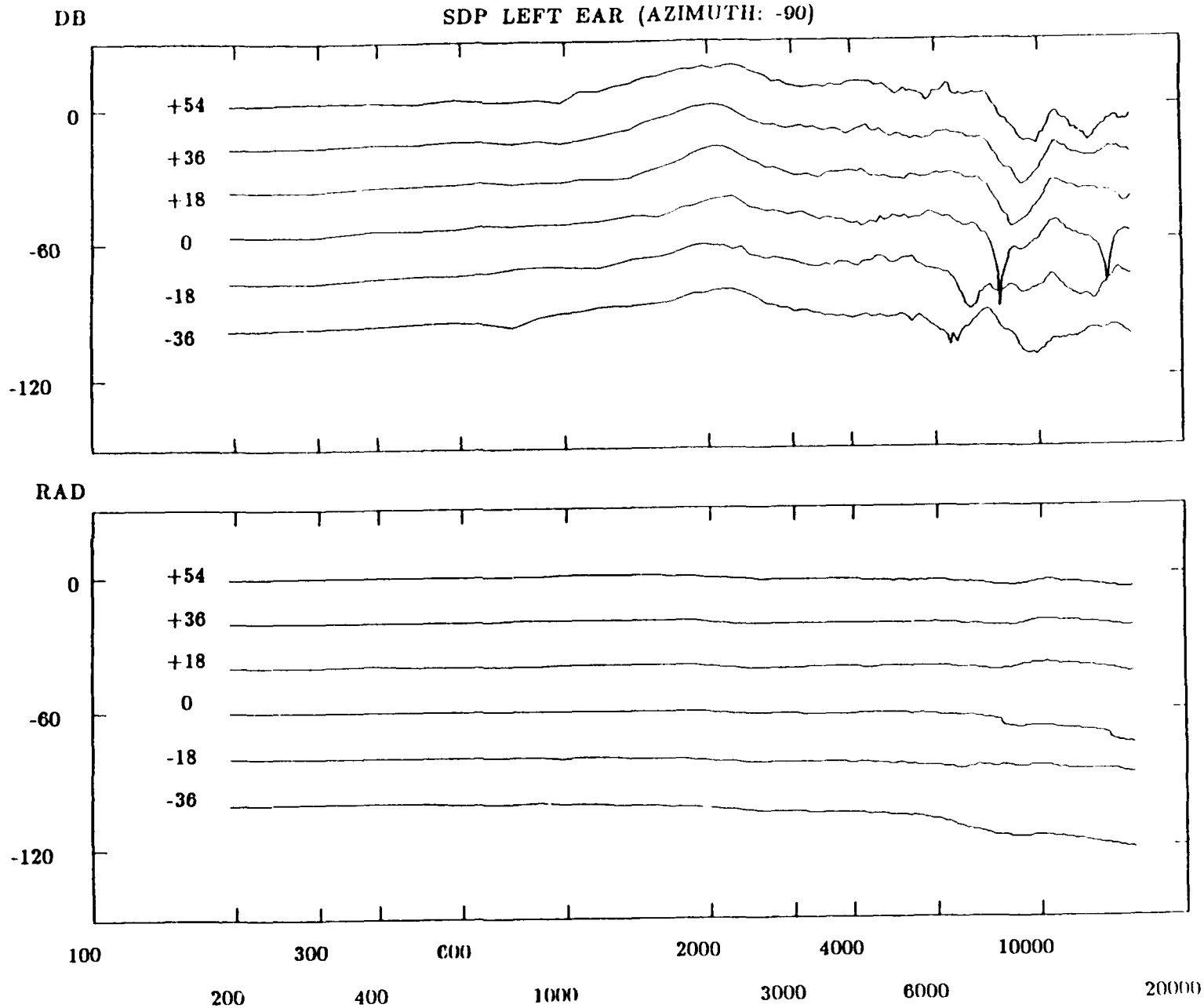
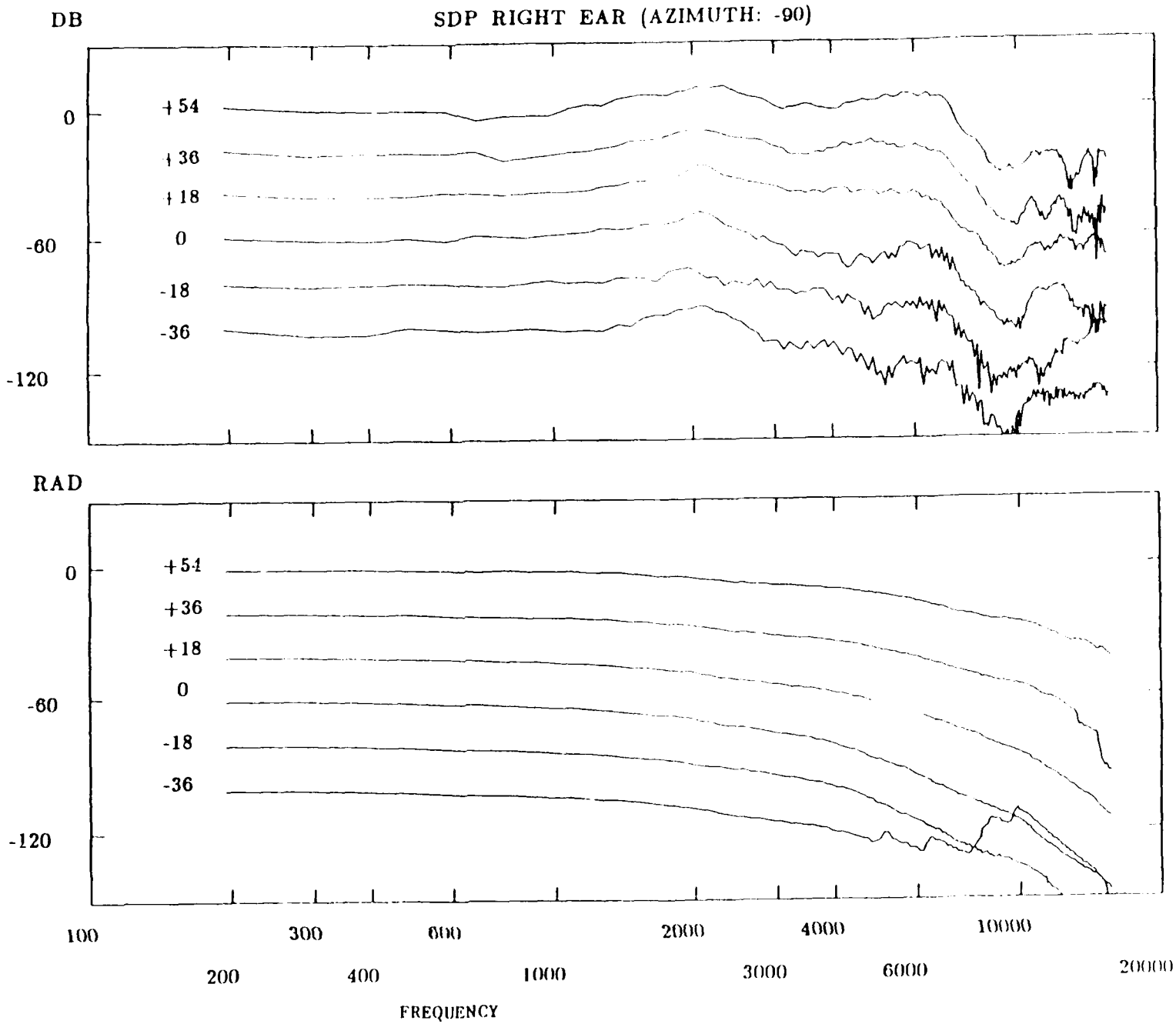


Figure 19. Same as Figure 17, except for a source at -90 degrees azimuth


SDP RIGHT EAR (AZIMUTH: -90)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

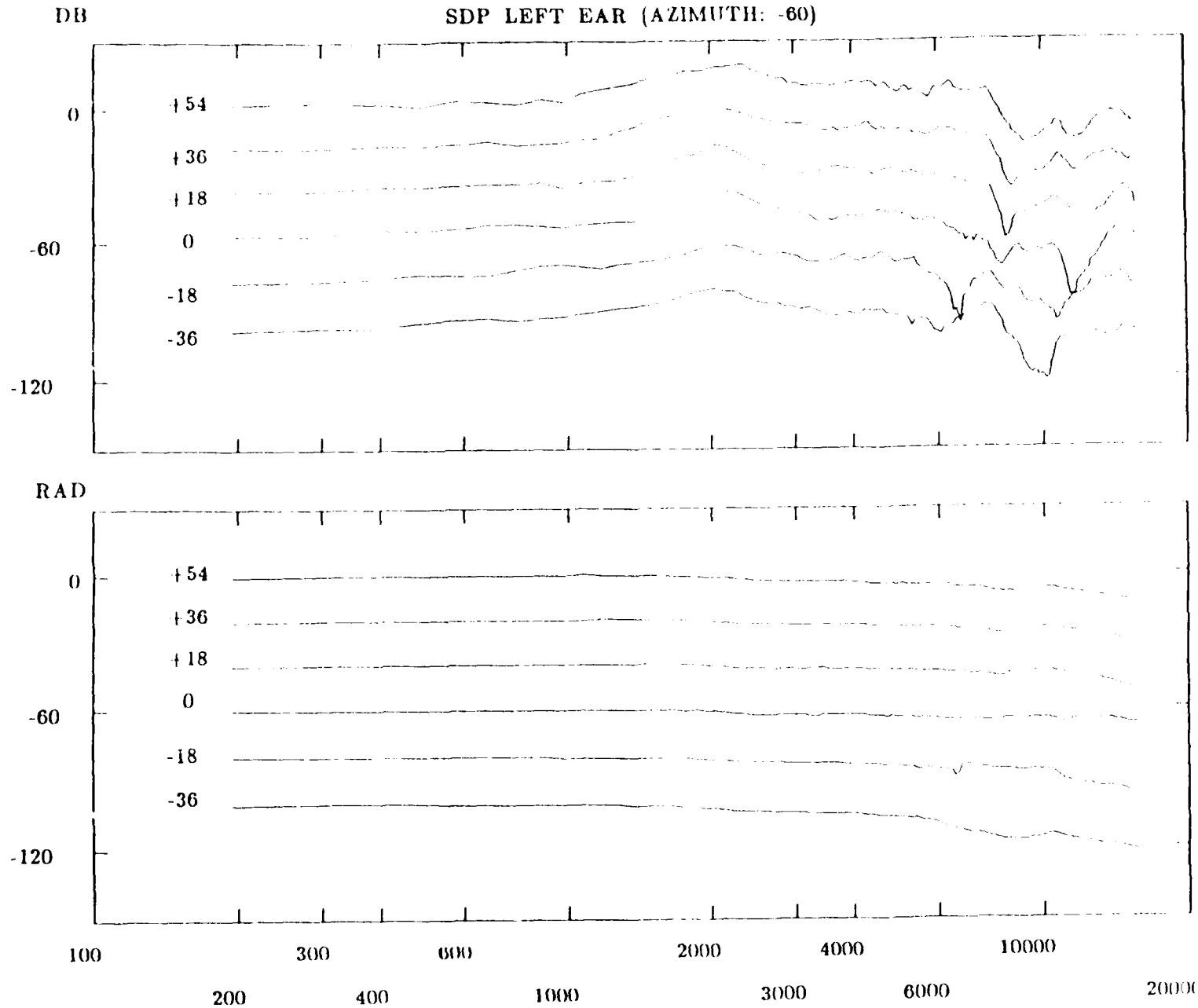
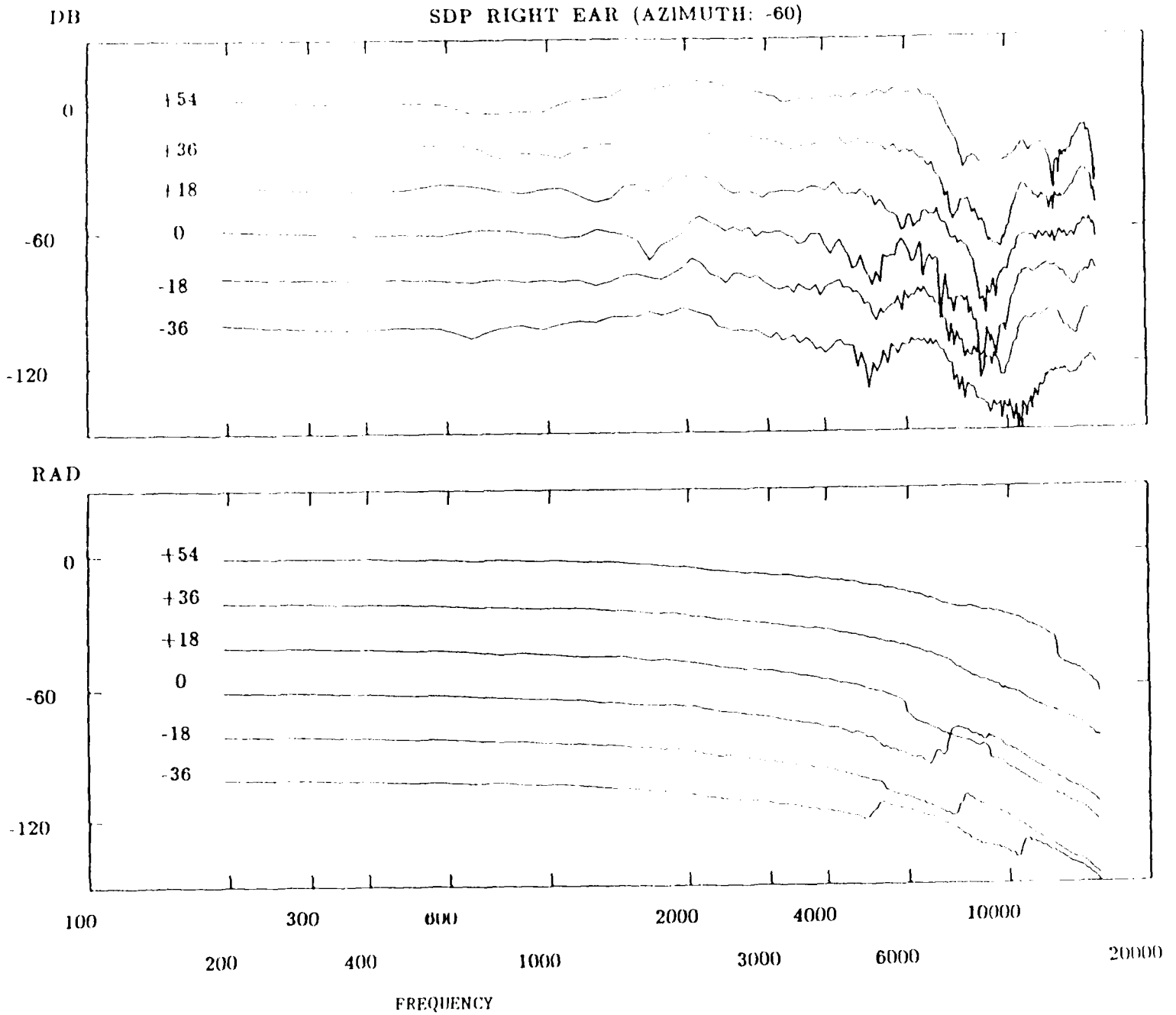


Figure 20. Same as Figure 17, except for a source at -60 degrees azimuth


SDP RIGHT EAR (AZIMUTH: -60)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

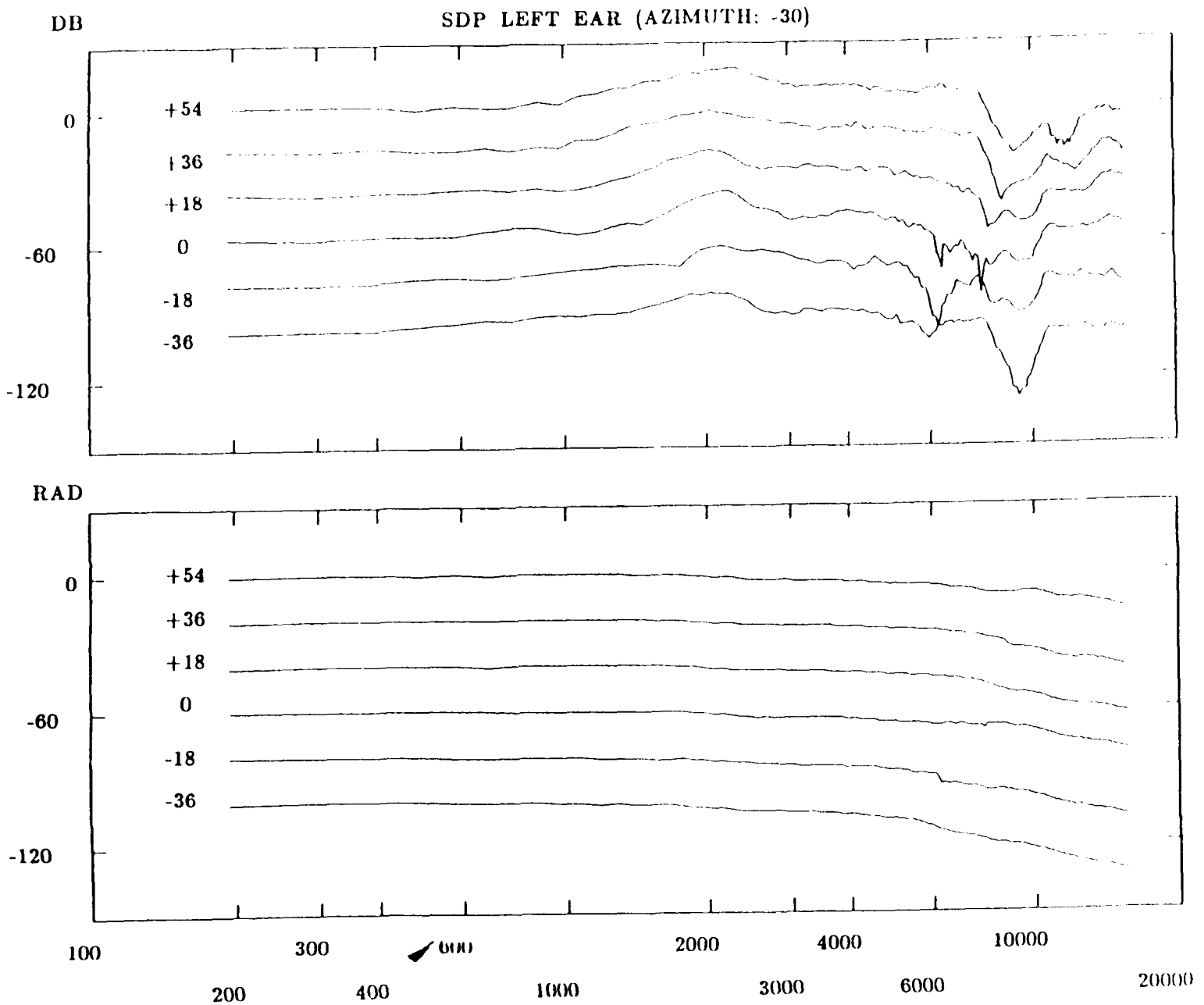
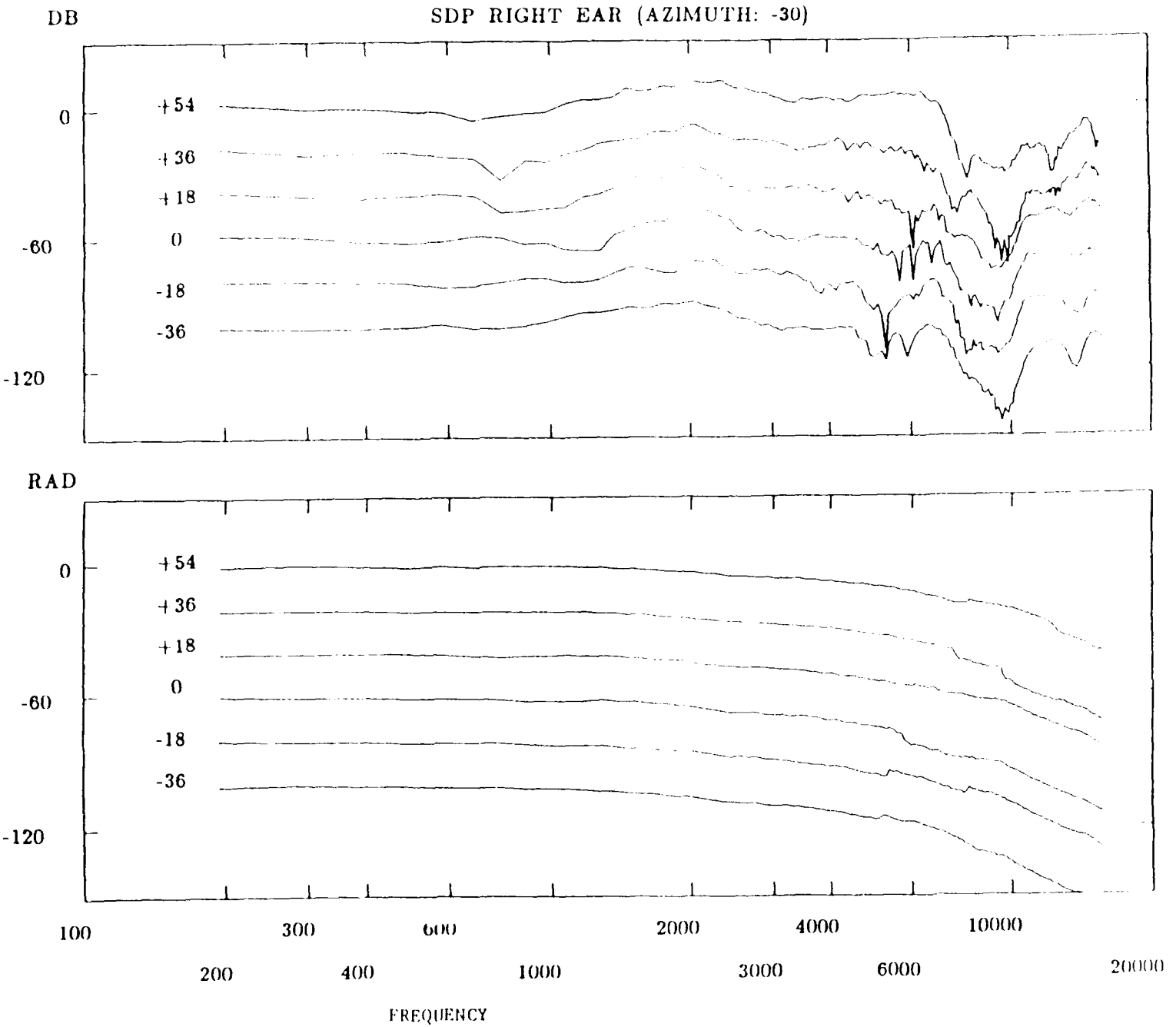


Figure 21. Same as Figure 17, except for a source at -30 degrees azimuth


SDP RIGHT EAR (AZIMUTH: -30)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

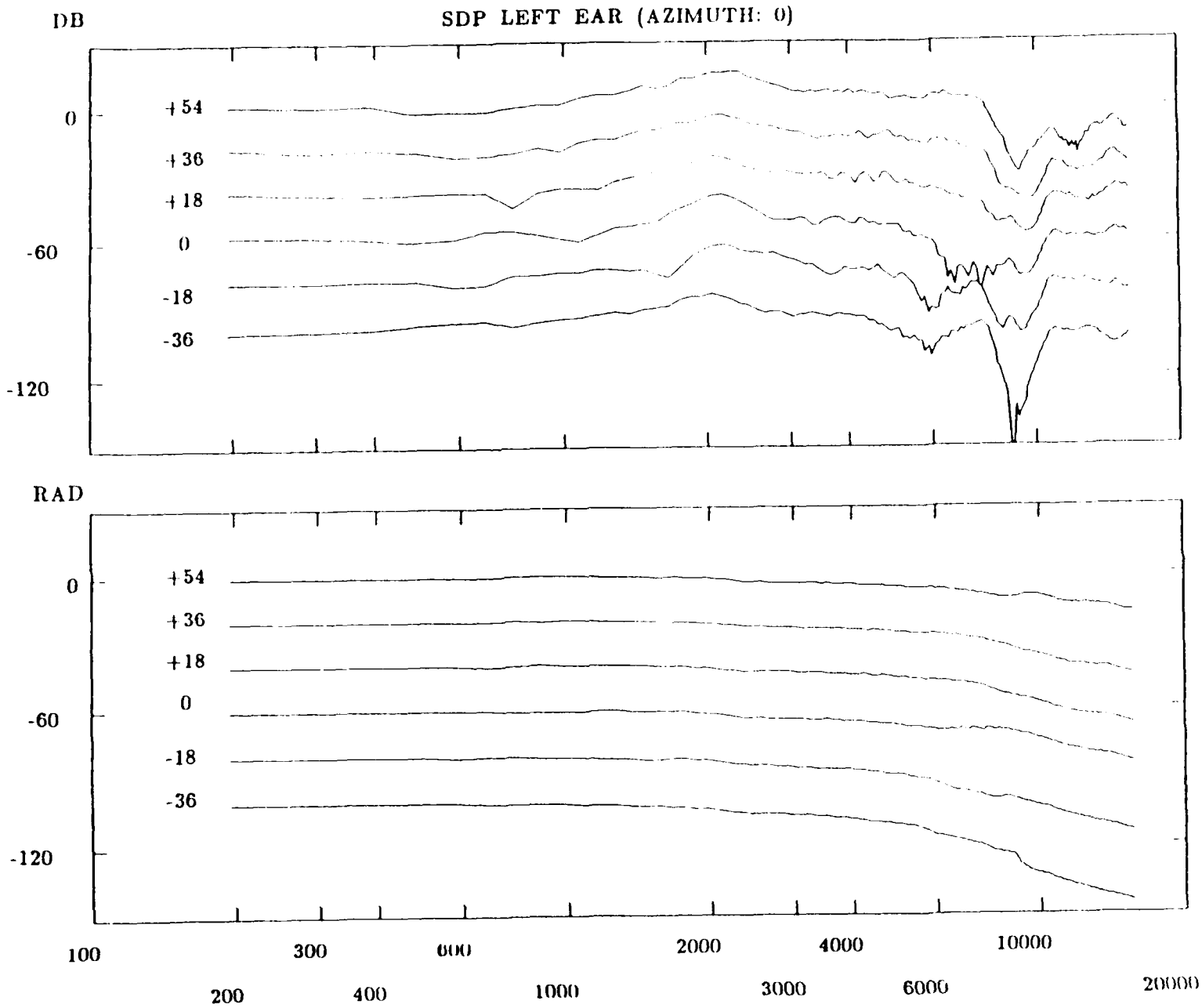
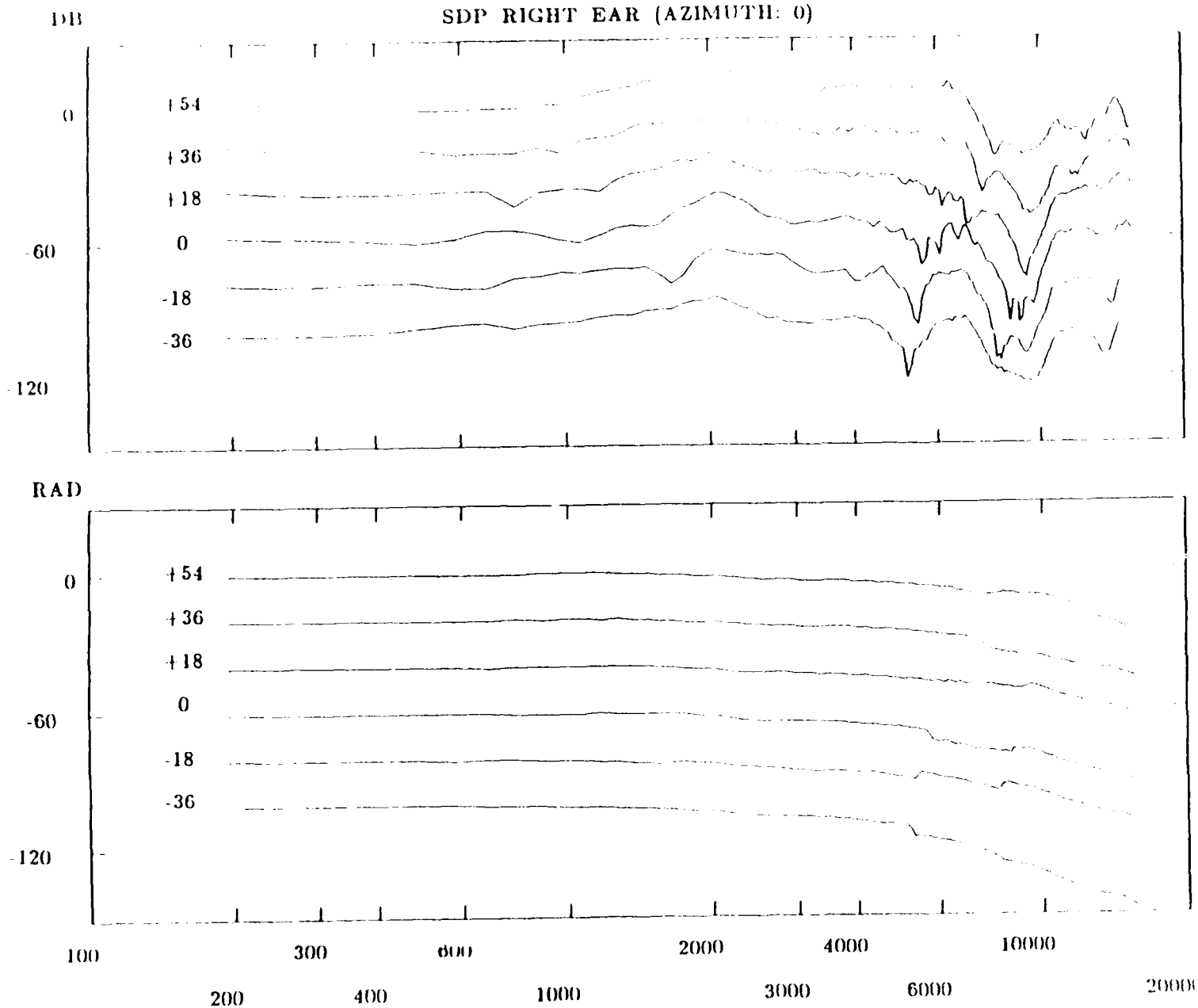


Figure 22. Same as Figure 17, except for a source at 0 degrees azimuth


SDP RIGHT EAR (AZIMUTH: 0)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
800
1000
2000
3000
4000
6000
10000
20000

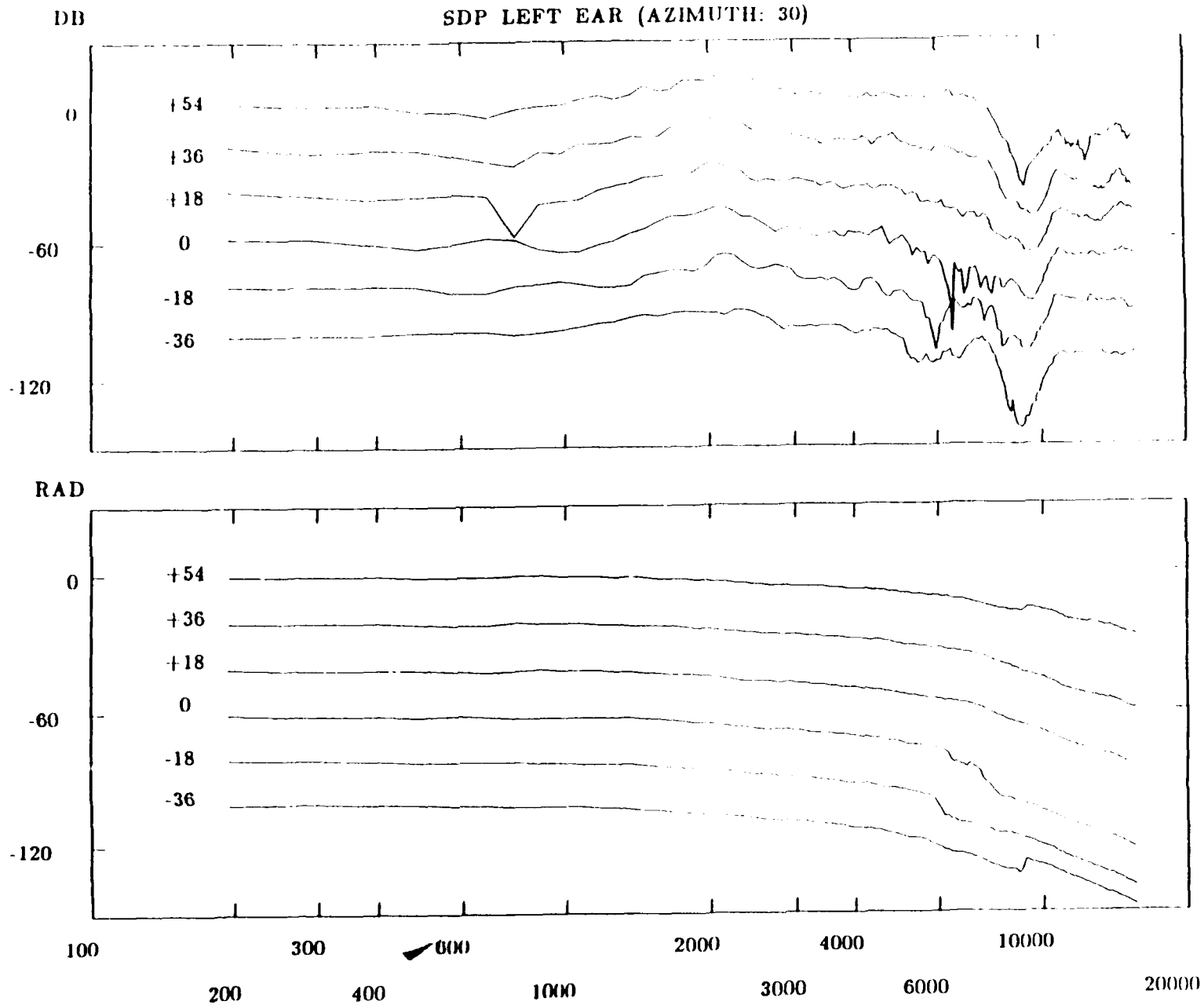
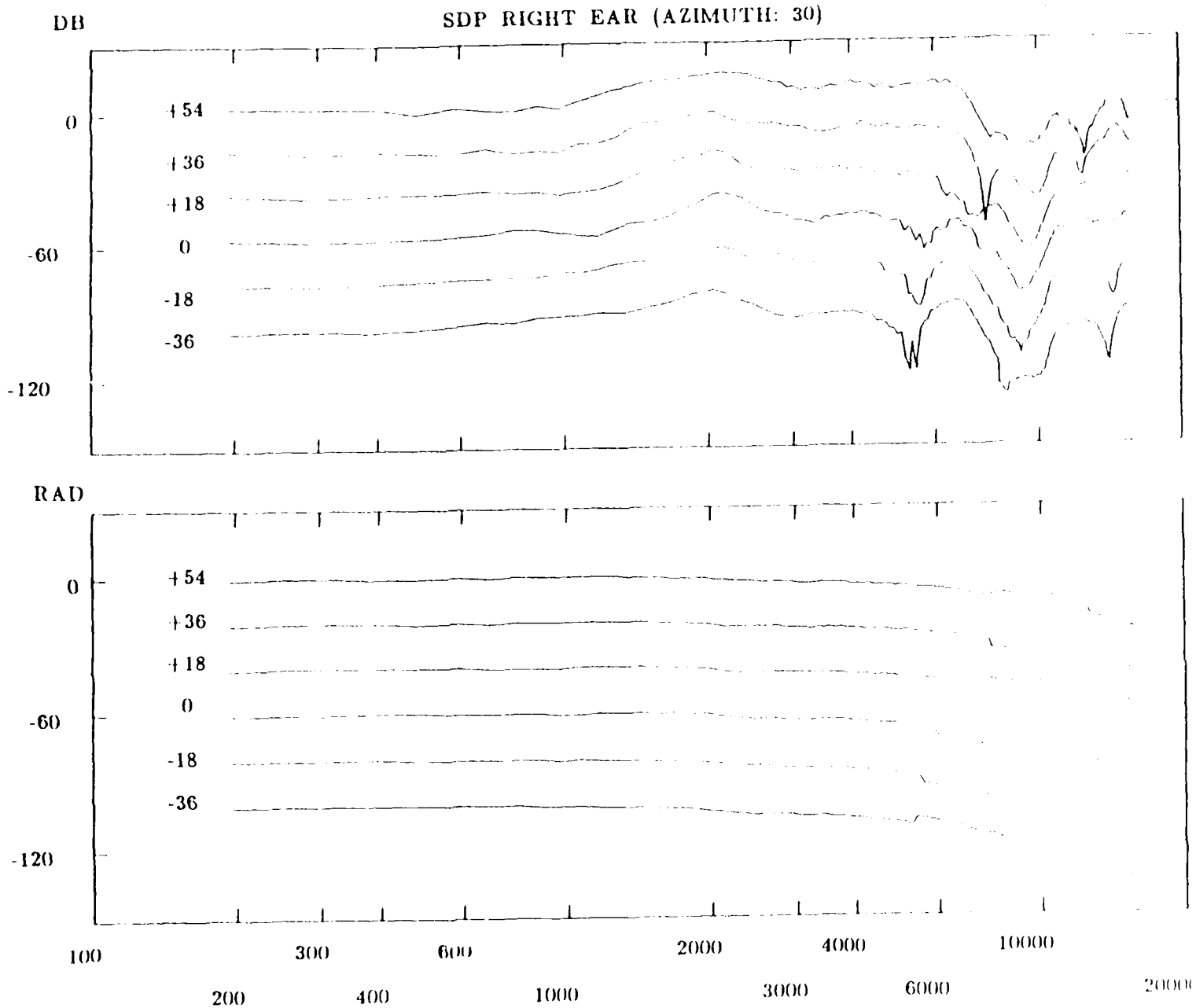


Figure 23. Same as Figure 17, except for a source at 30 degrees azimuth


SDP RIGHT EAR (AZIMUTH: 30)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000

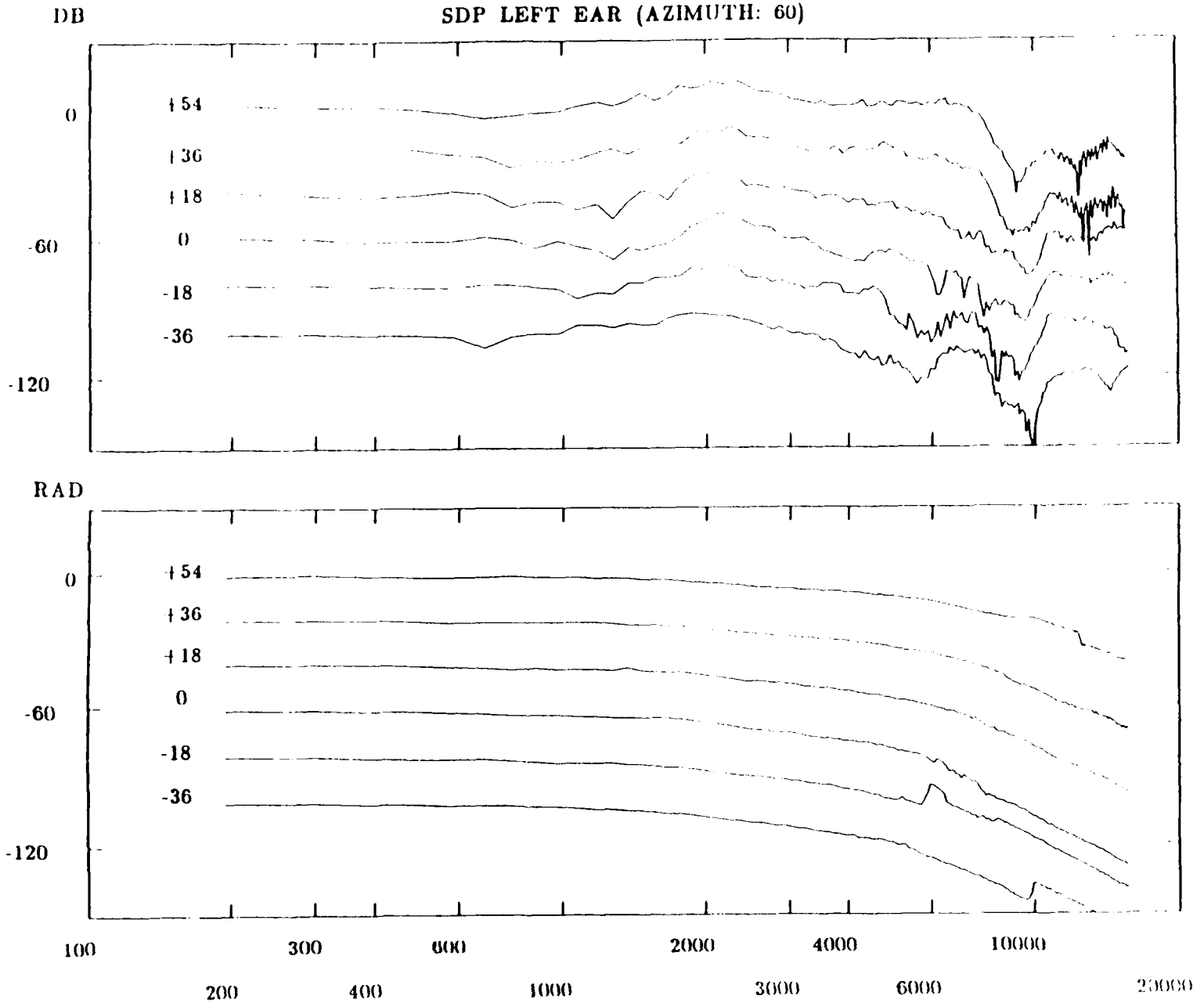
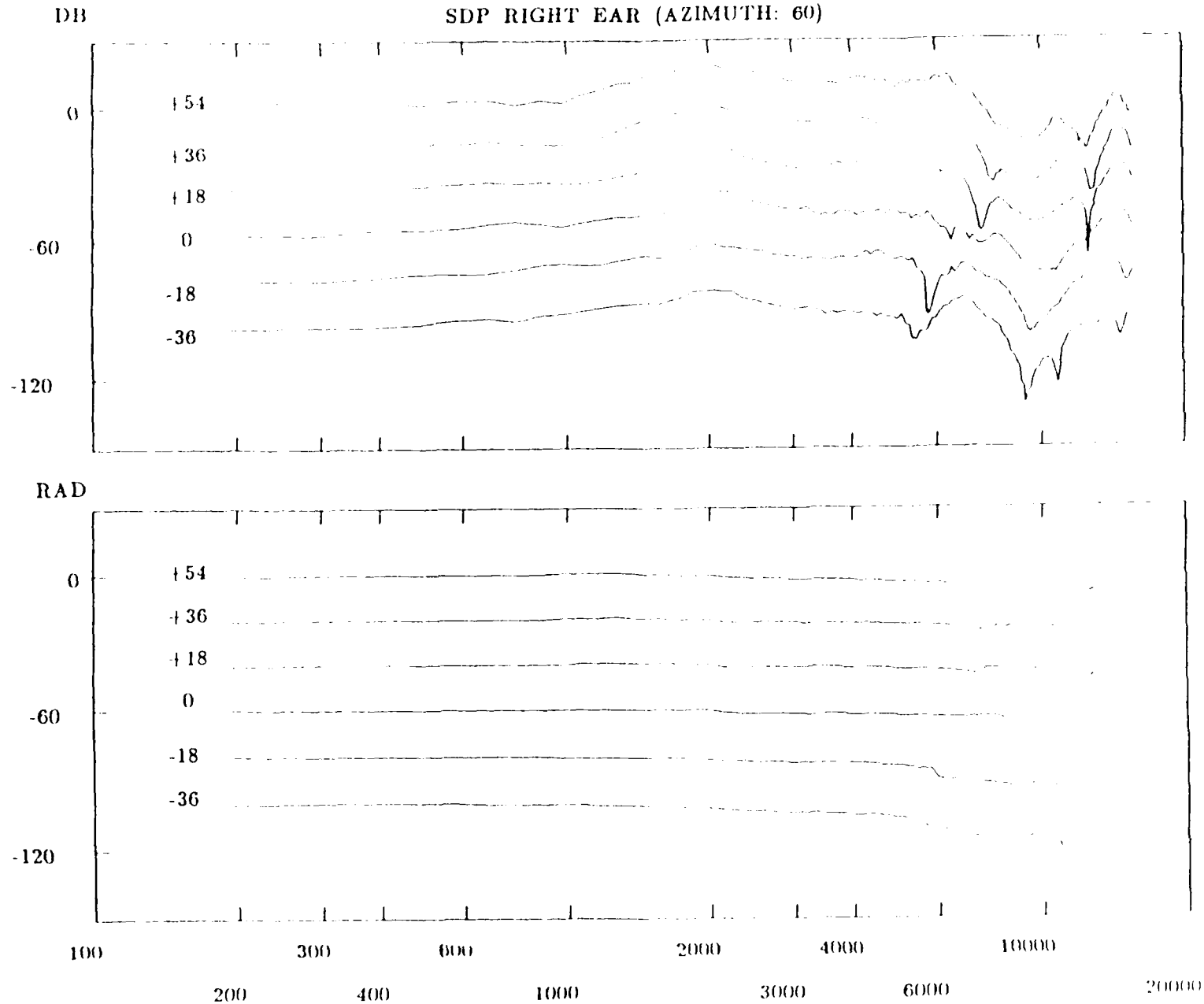


Figure 24. Same as Figure 17, except for a source at 60 degrees azimuth


DB
SDP RIGHT EAR (AZIMUTH: 60)
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000

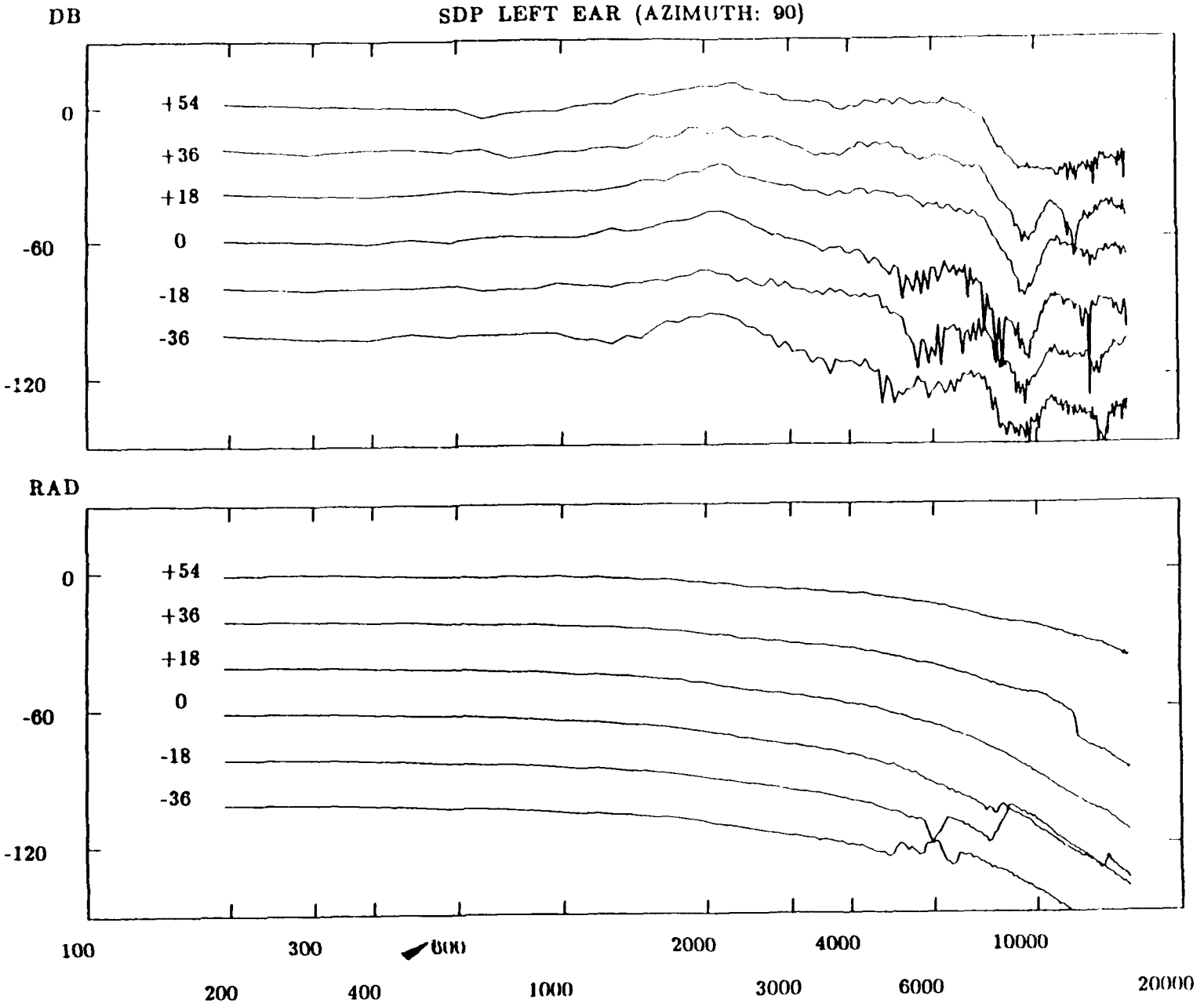
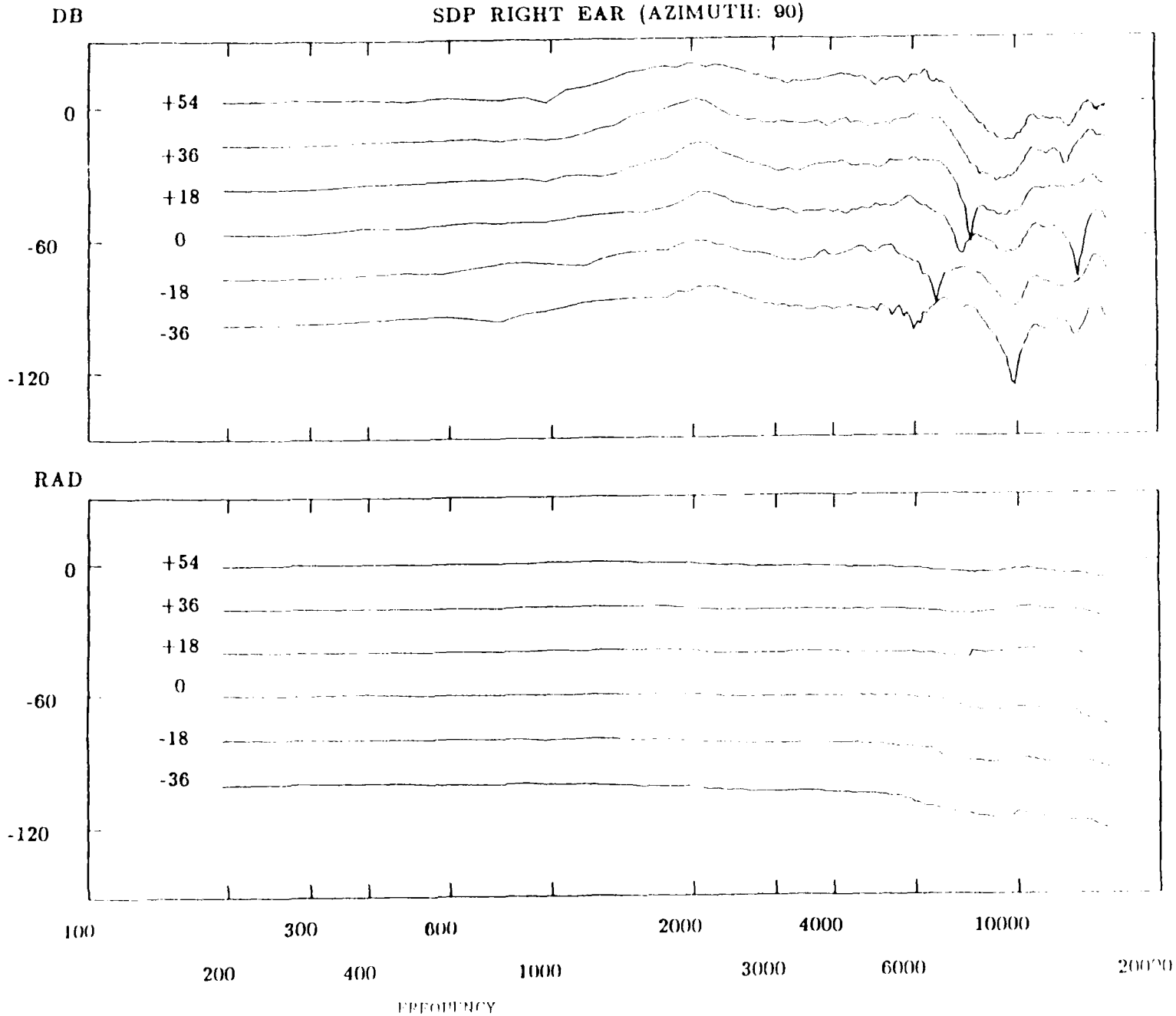


Figure 25. Same as Figure 17, except for a source at 90 degrees azimuth


DB
SDP RIGHT EAR (AZIMUTH: 90)
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

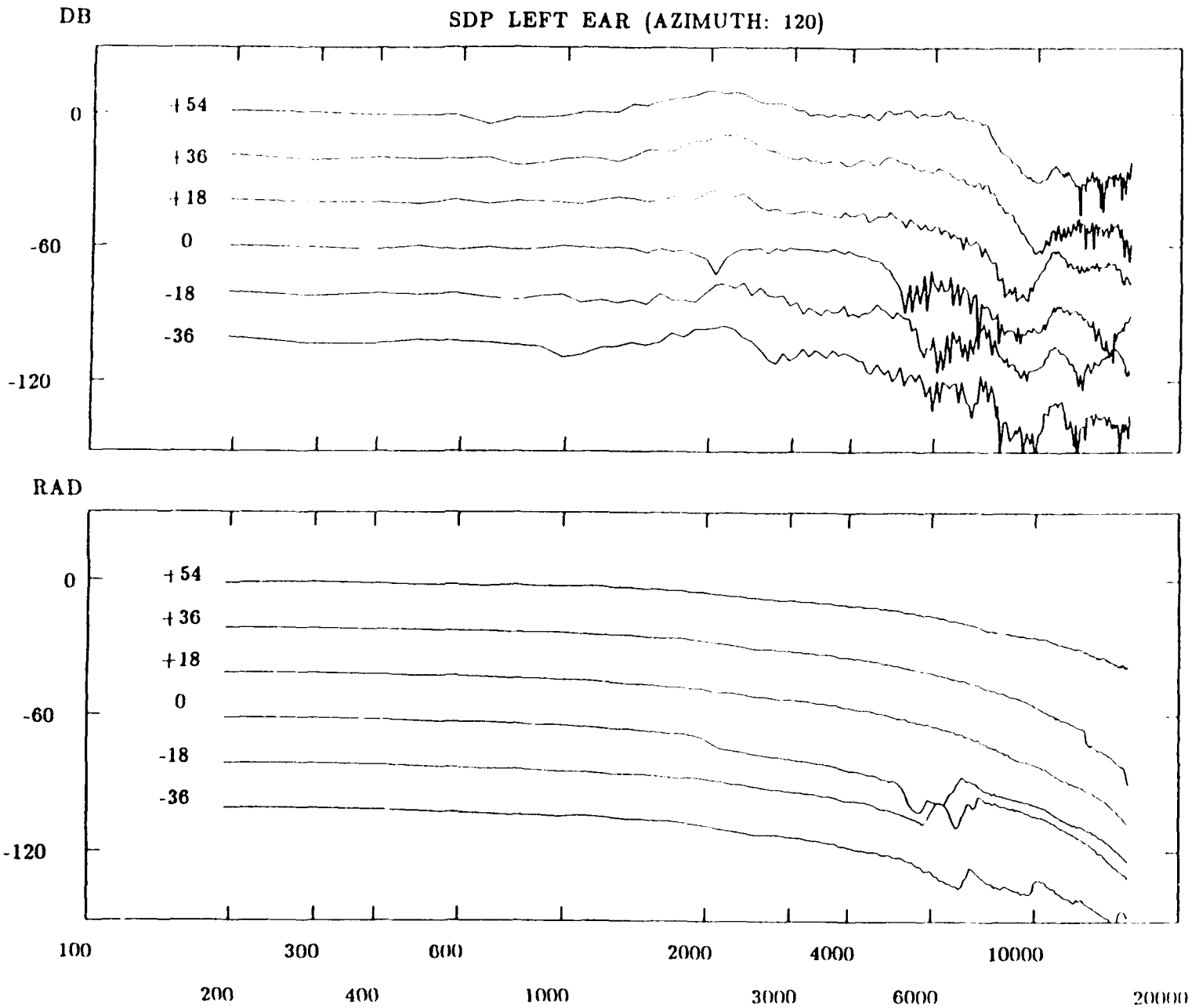
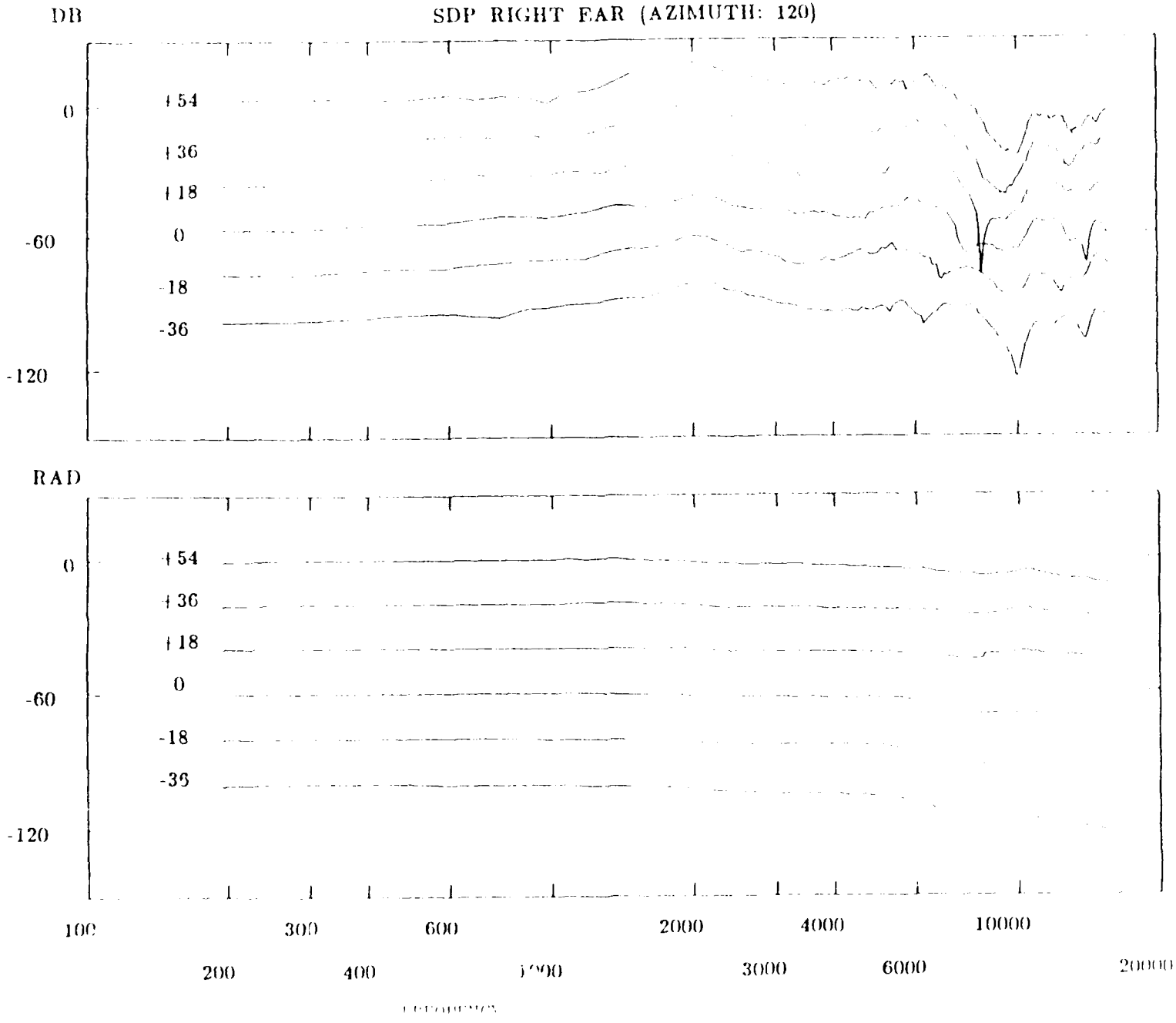


Figure 26. Same as Figure 17, except for a source at 120 degrees azimuth


DB
SDP RIGHT EAR (AZIMUTH: 120)
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
2000
3000
4000
6000
10000
20000

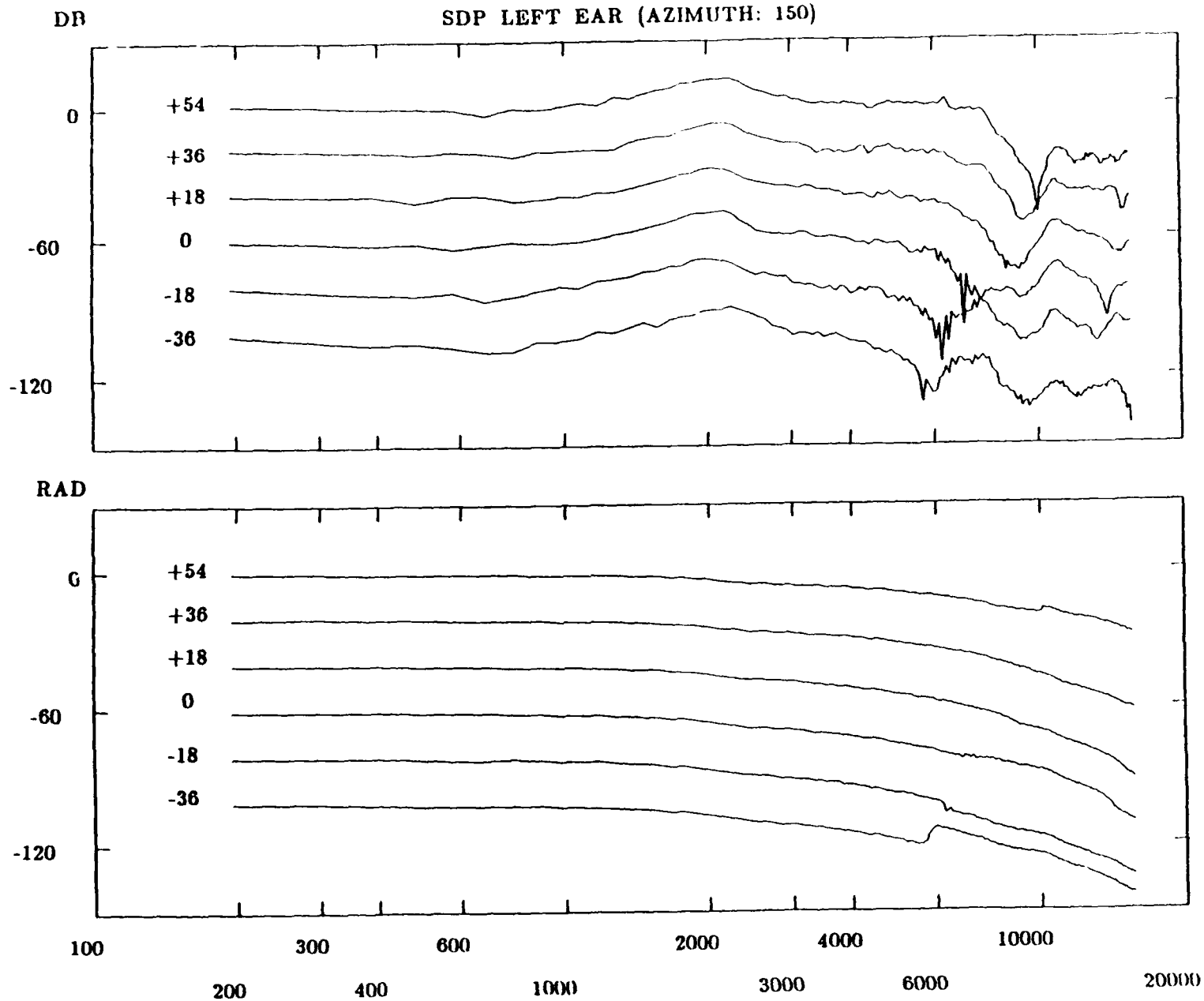
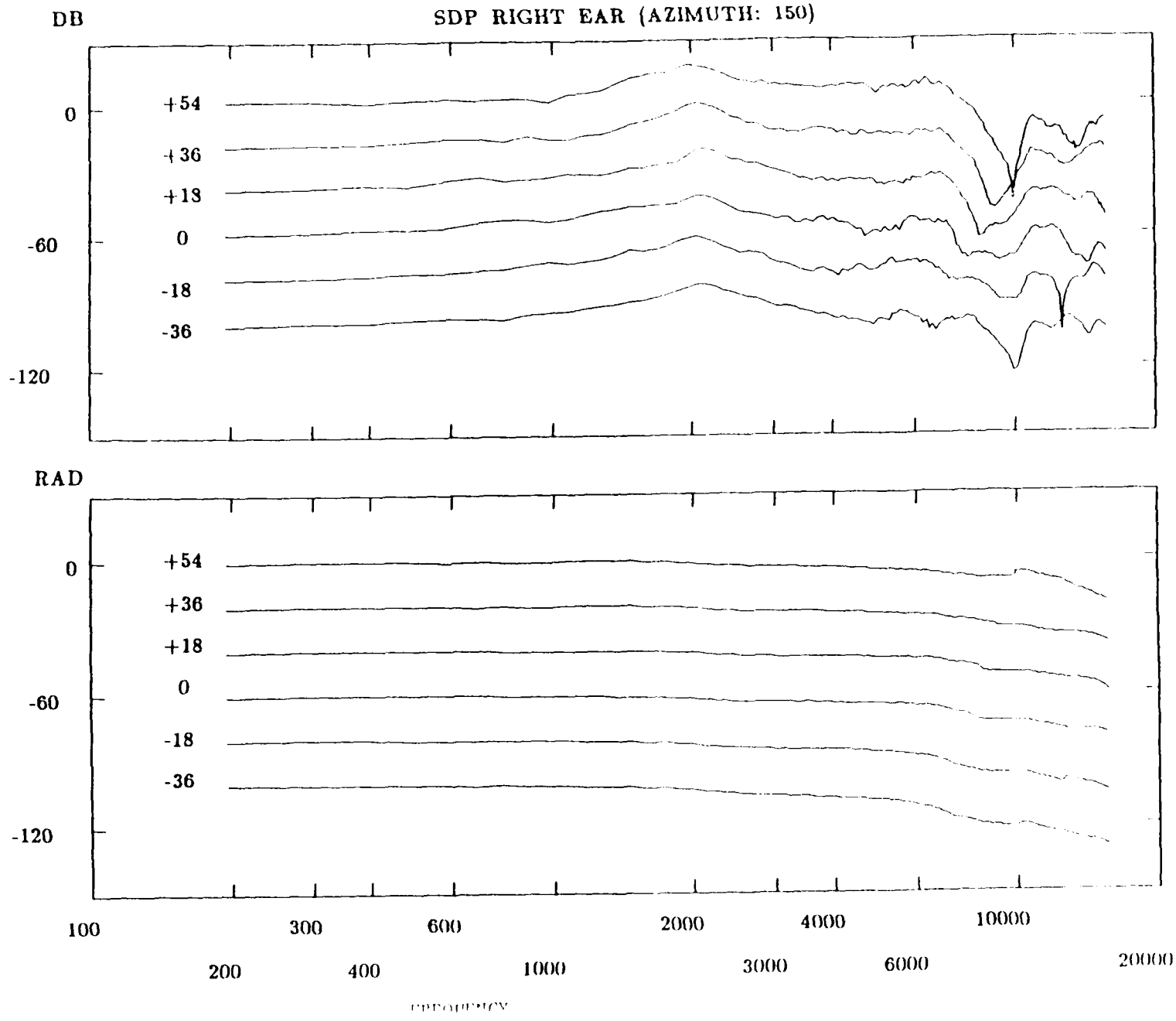


Figure 27. Same as Figure 17, except for a source at 150 degrees azimuth


SDP RIGHT EAR (AZIMUTH: 150)
DB
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

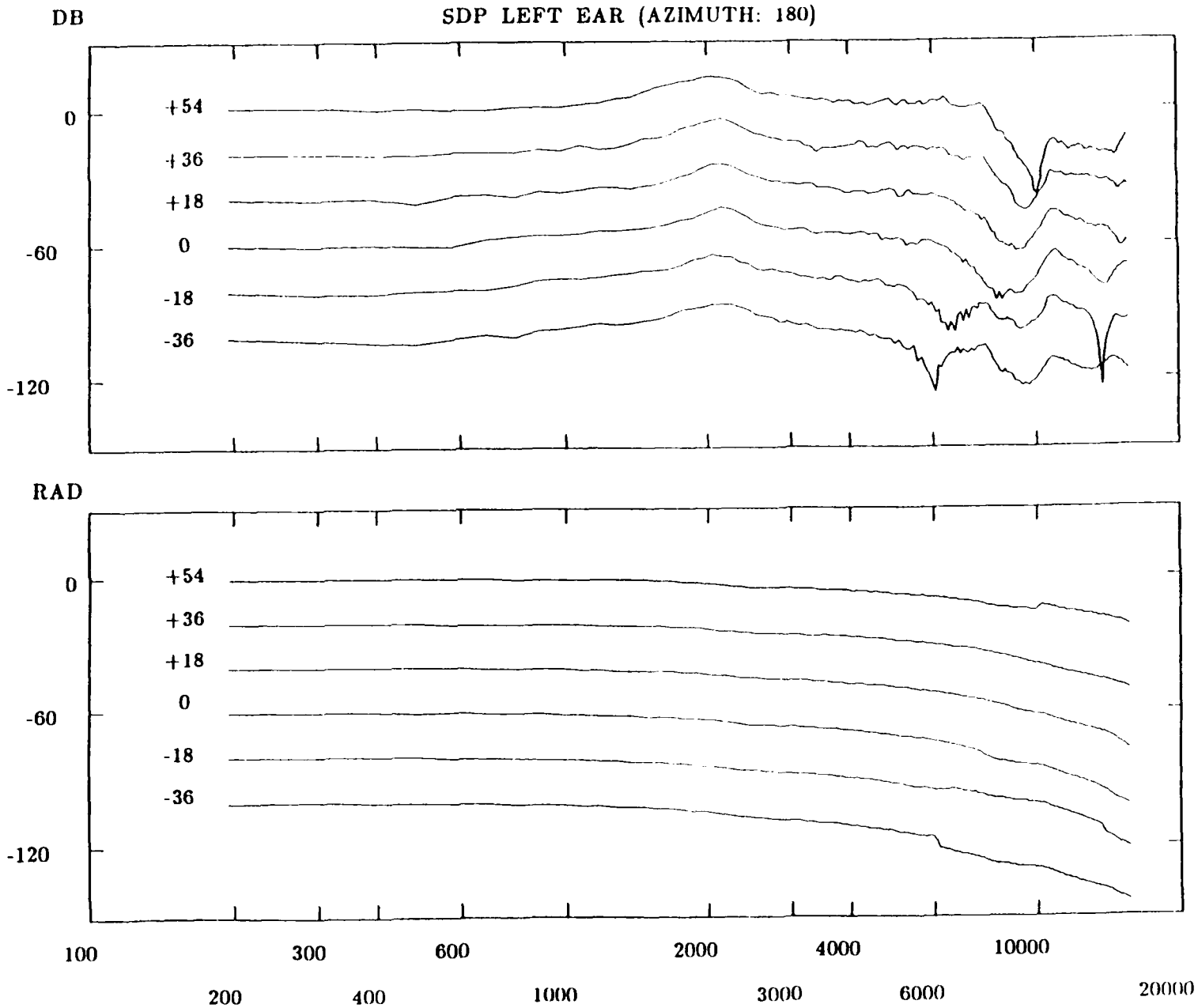
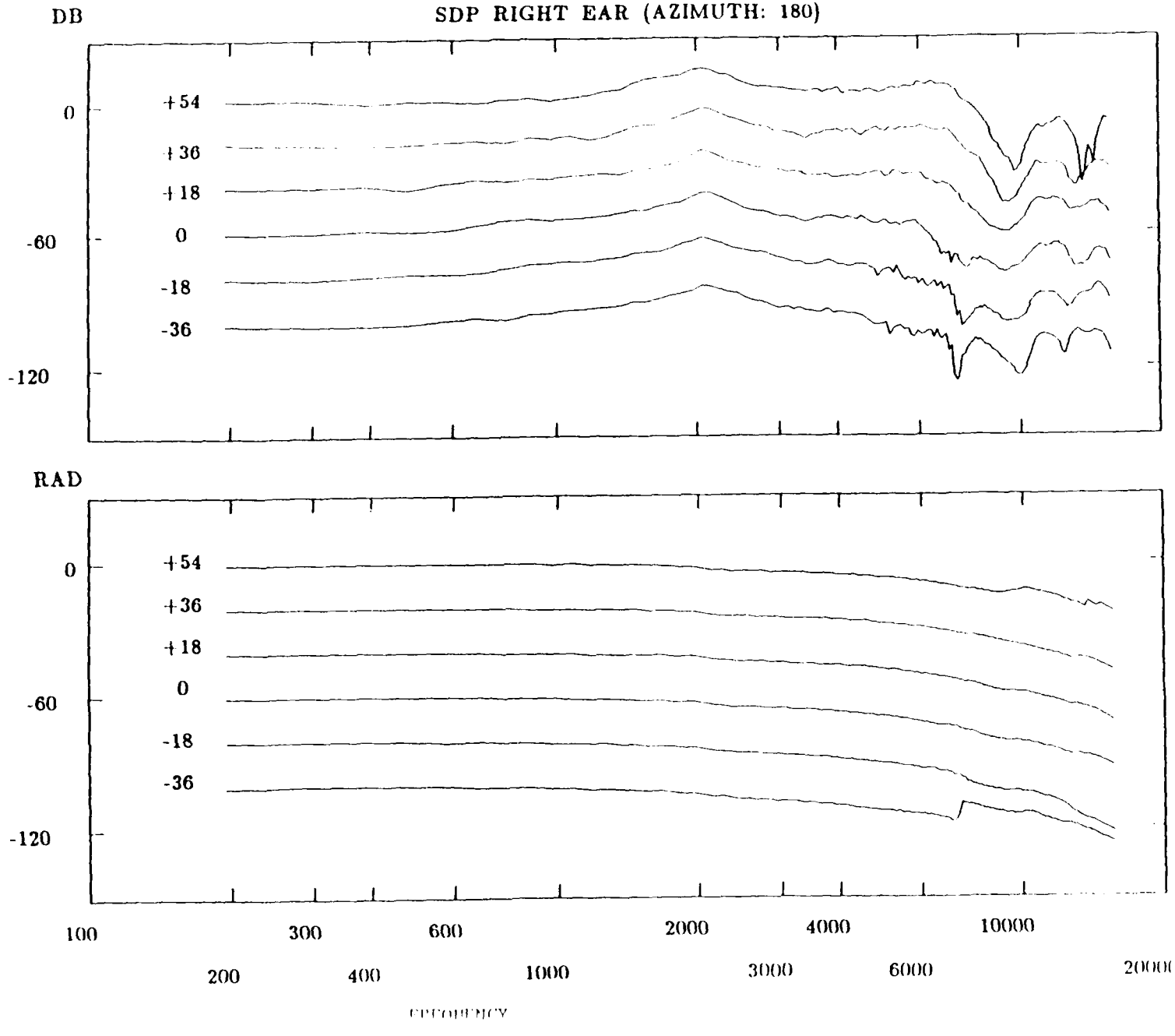


Figure 28. Same as Figure 17, except for a source at 180 degrees azimuth


DB
SDP RIGHT EAR (AZIMUTH: 180)
0
-60
-120
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

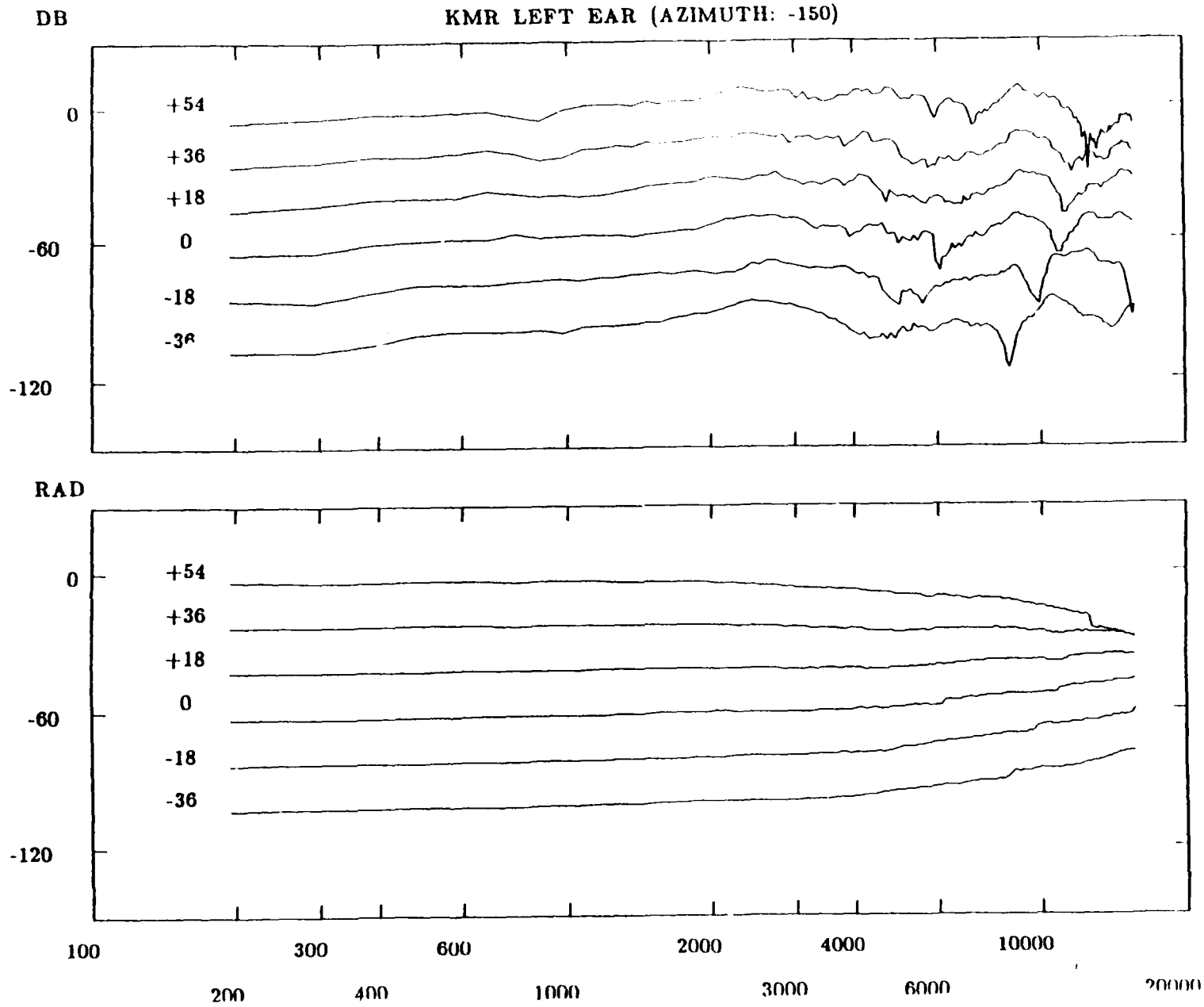


Figure 29. HRTF measurements from KEMAR for a source at -150 degrees azimuth (i.e., 150 degrees on the mannequin's left side) and +54, +36, +18, 0, -18, -36 degrees elevation. The left ear magnitude function is plotted in the upper panel in dB coordinates and the left ear phase function is plotted in the lower panel in radian coordinates. The functions for the six elevations are displaced vertically by 20 dB for visibility. The measurements were obtained using a periodic pseudorandom noise

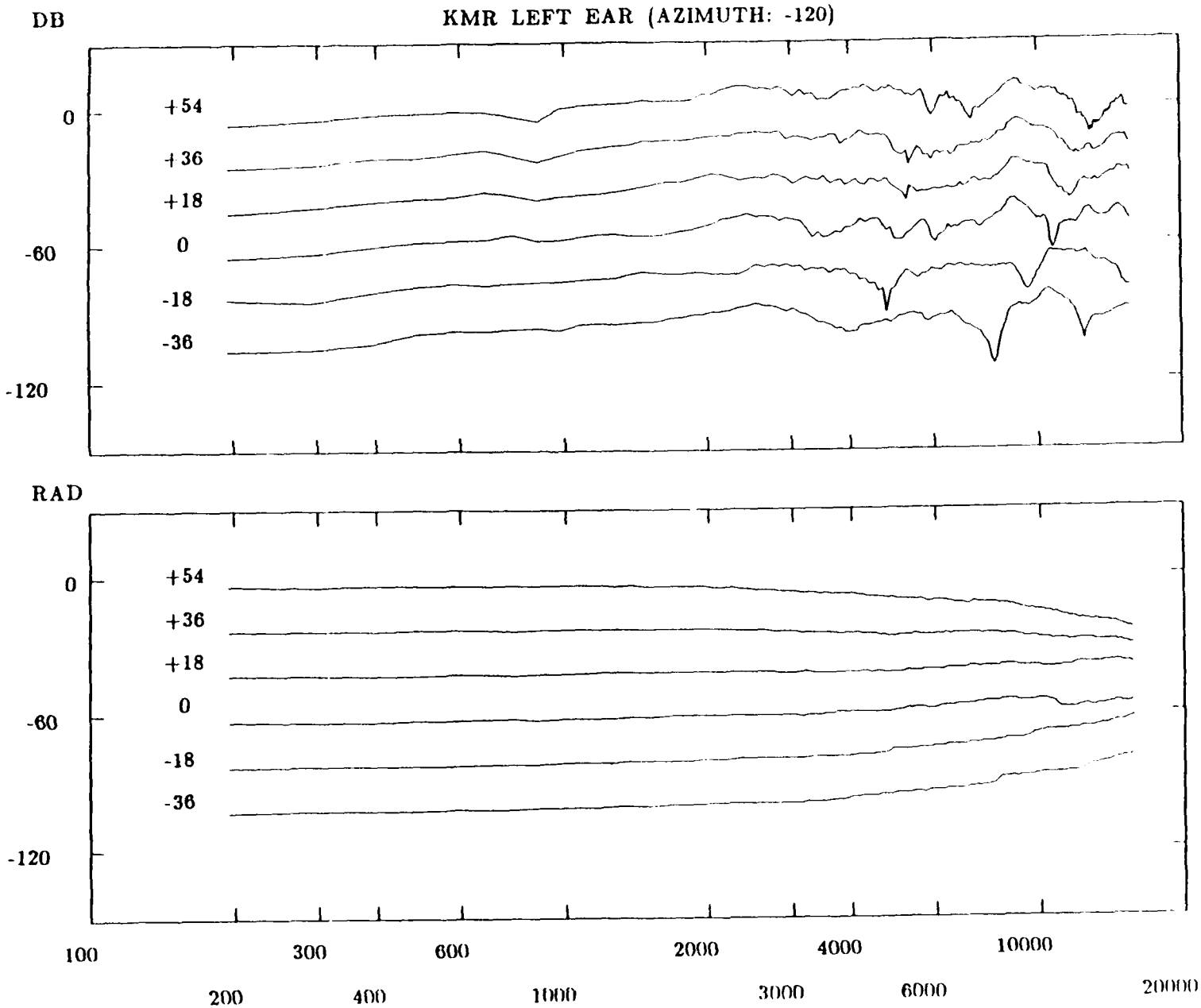


Figure 30. Same as Figure 29, except for a source at -120 degrees azimuth

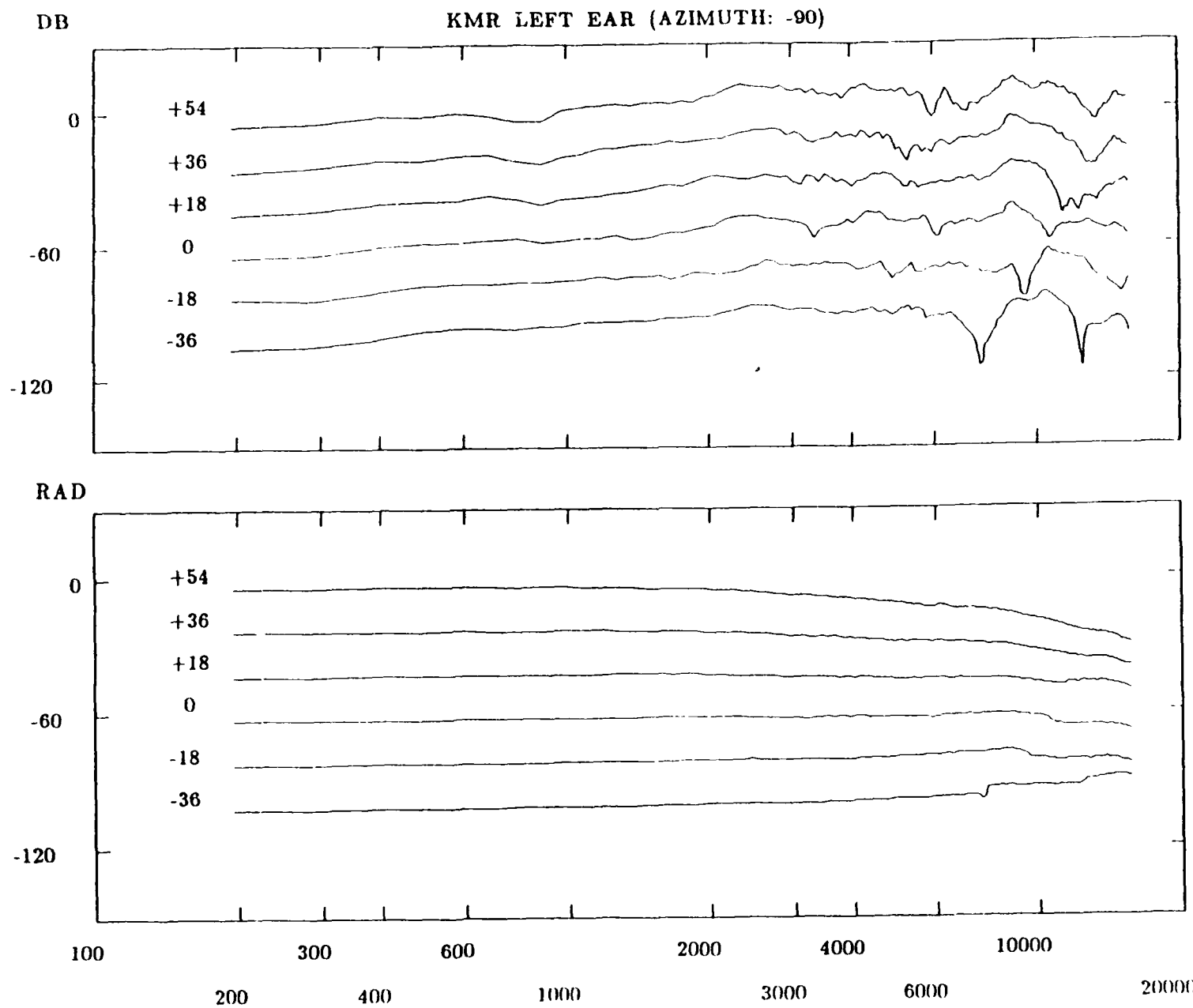


Figure 31. Same as Figure 29, except for a source at -90 degrees azimuth

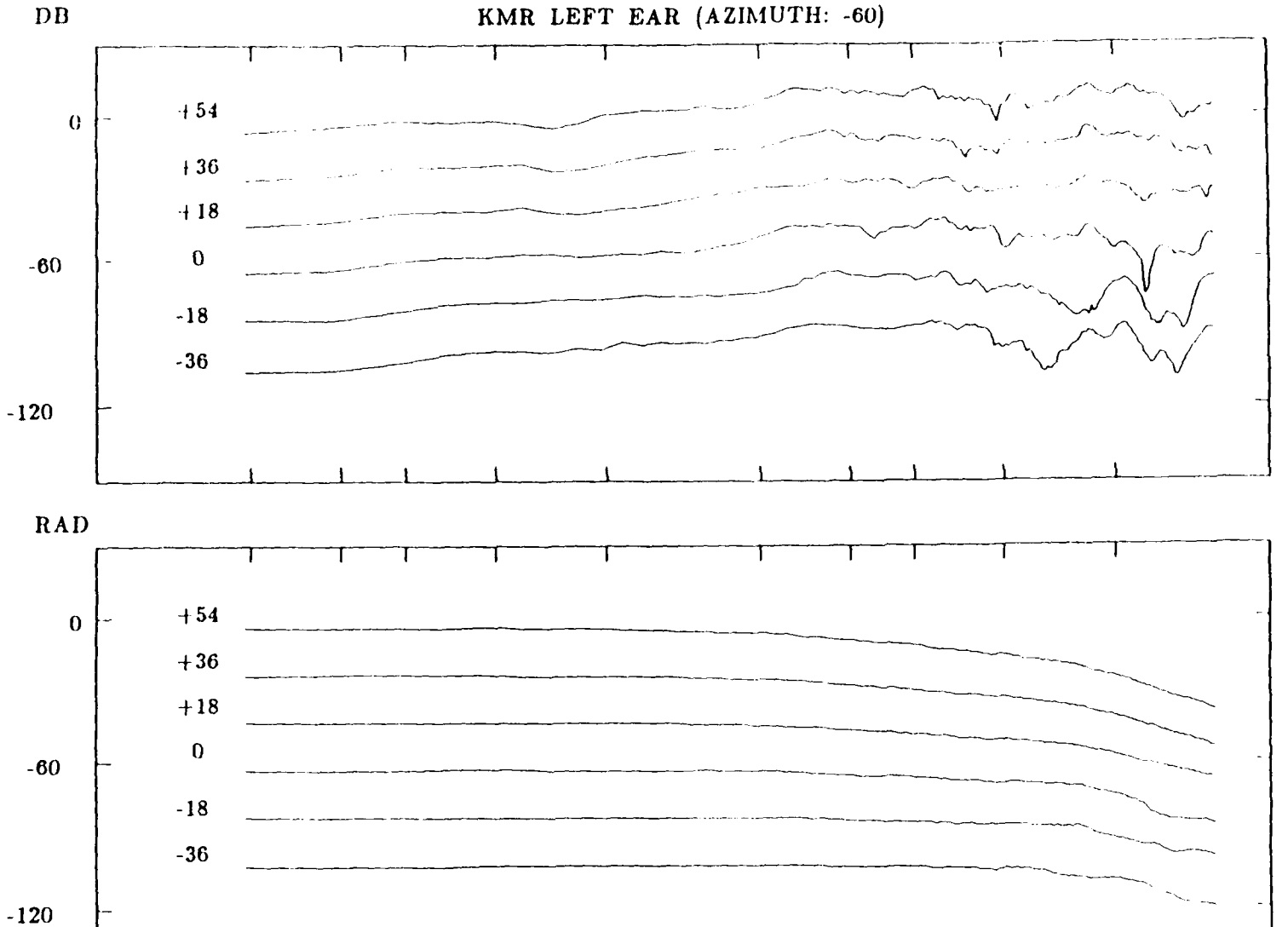


Figure 32. Same as Figure 29, except for a source at -60 degrees azimuth

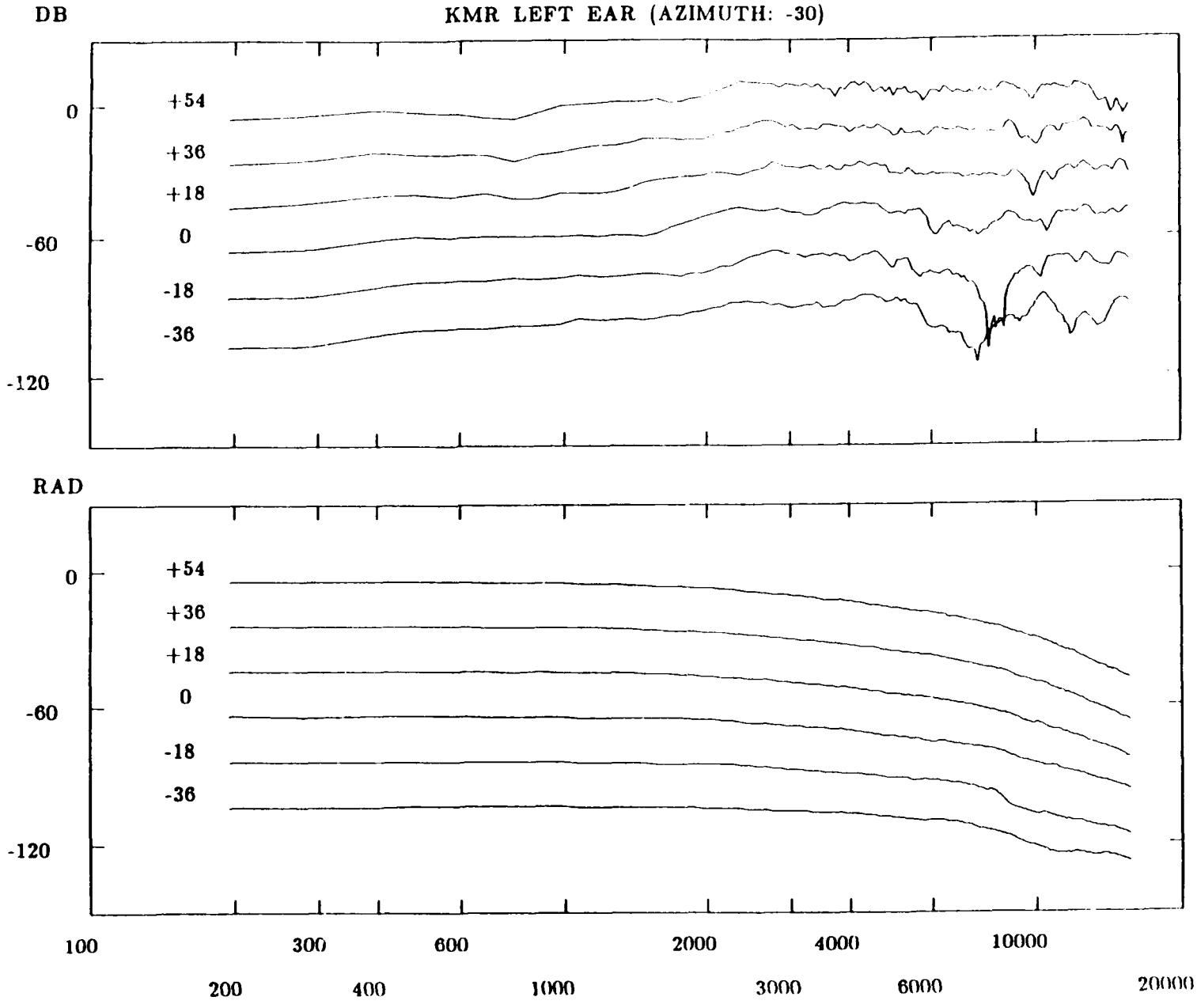


Figure 33. Same as Figure 29, except for a source at -30 degrees azimuth

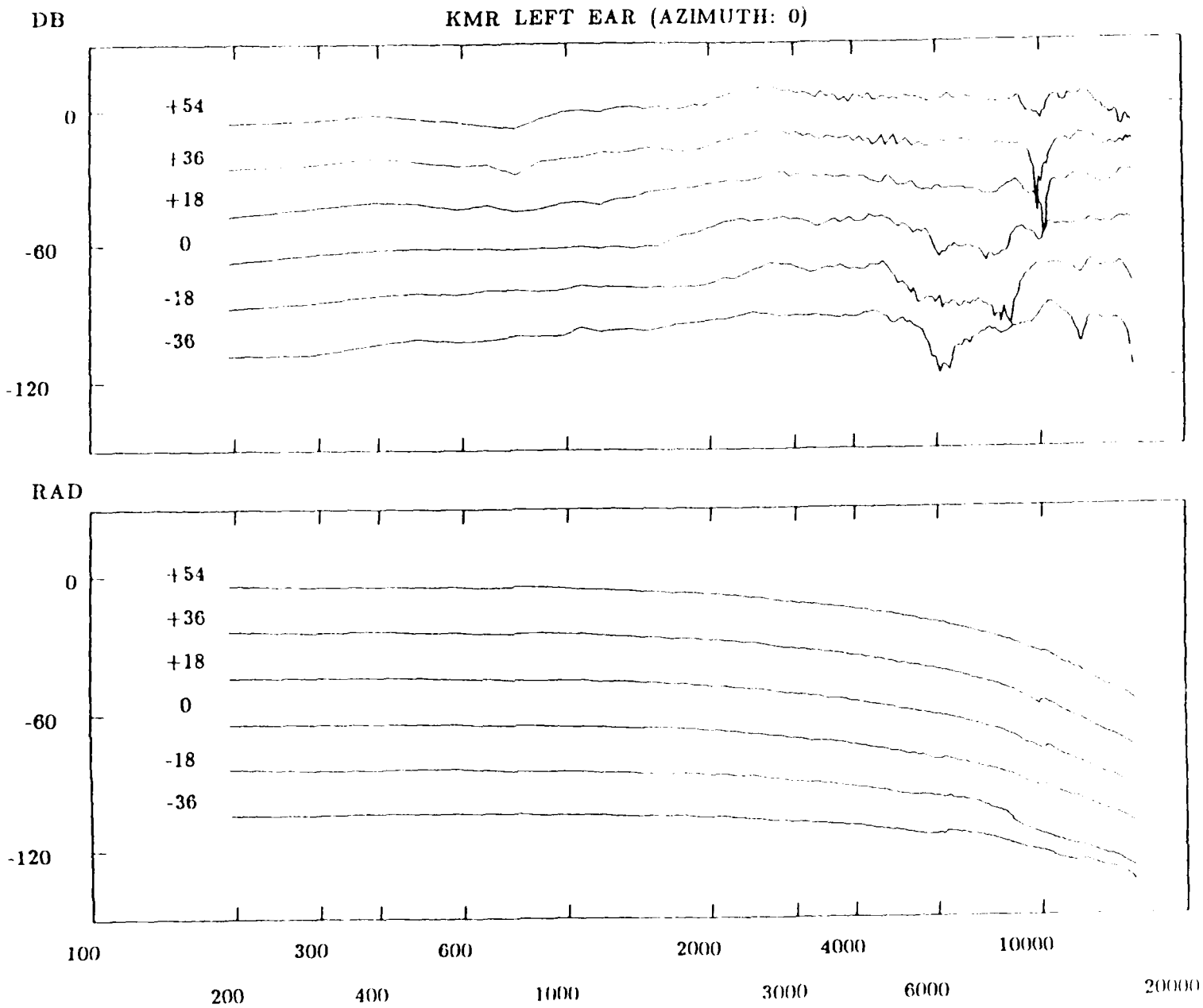


Figure 34. Same as Figure 29, except for a source at 0 degrees azimuth

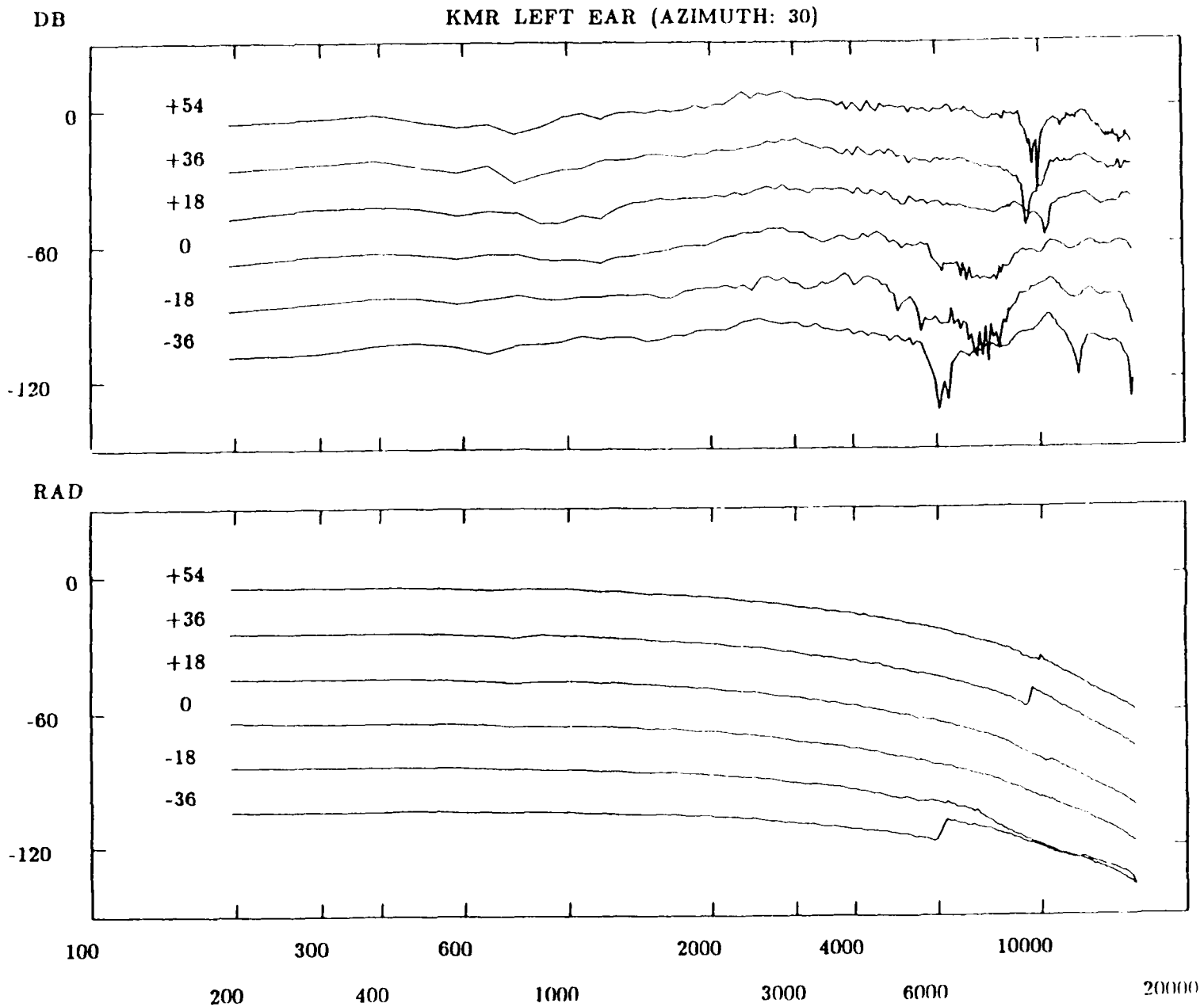


Figure 35. Same as Figure 29, except for a source at 30 degrees azimuth

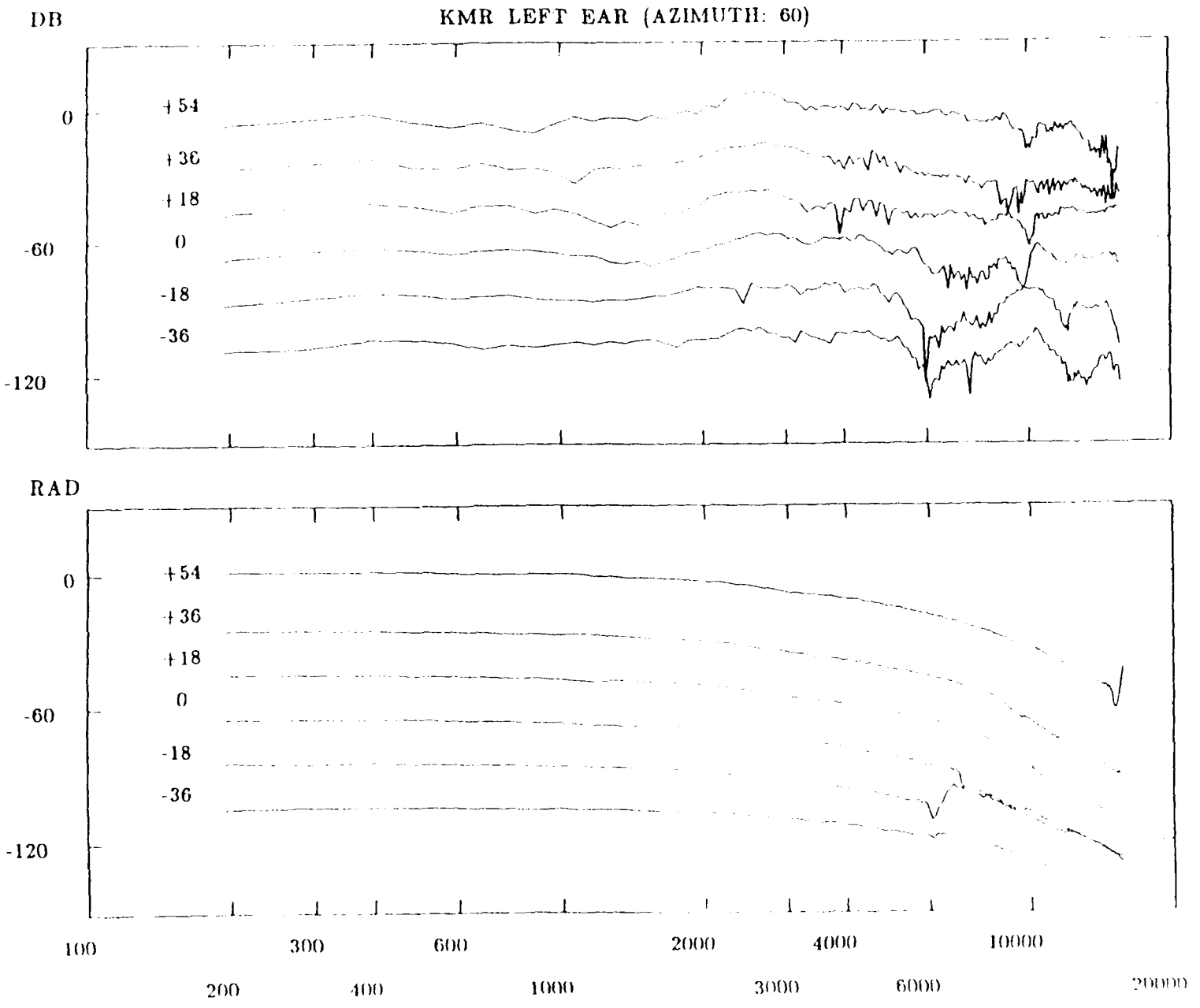


Figure 36. Same as Figure 29, except for a source at 60 degrees azimuth

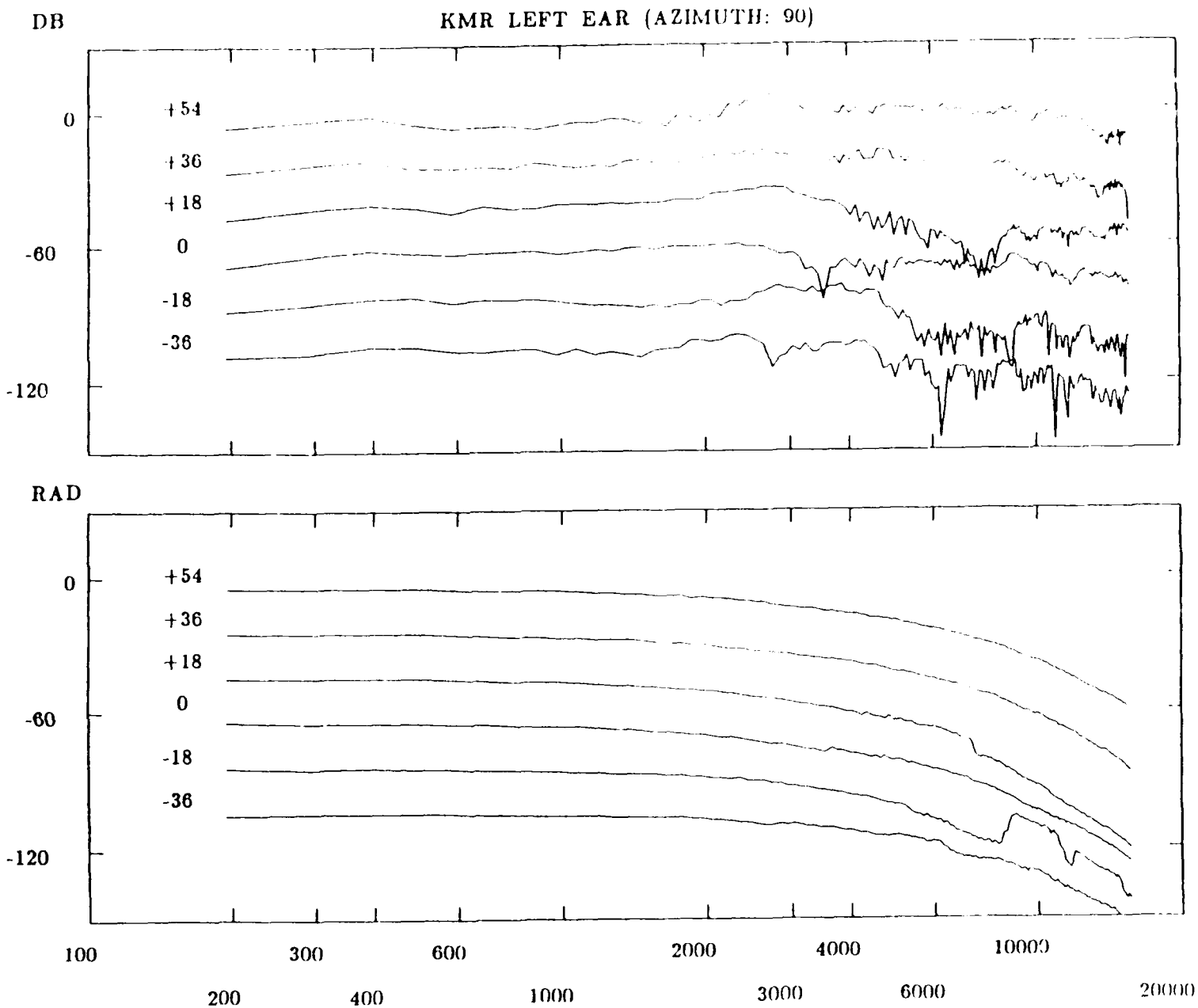


Figure 37. Same as Figure 29, except for a source at 90 degrees azimuth

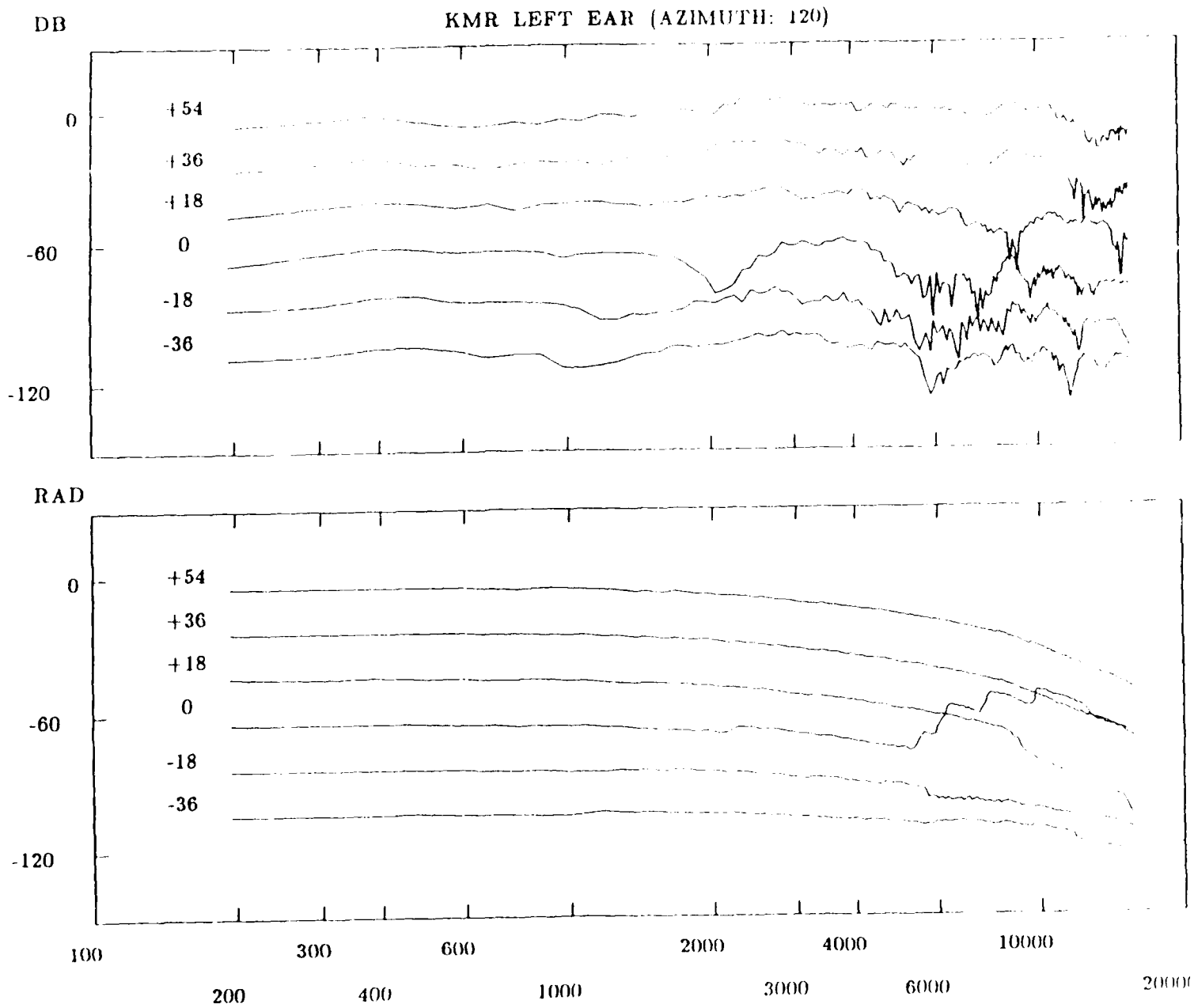


Figure 38. Same as Figure 29, except for a source at 120 degrees azimuth

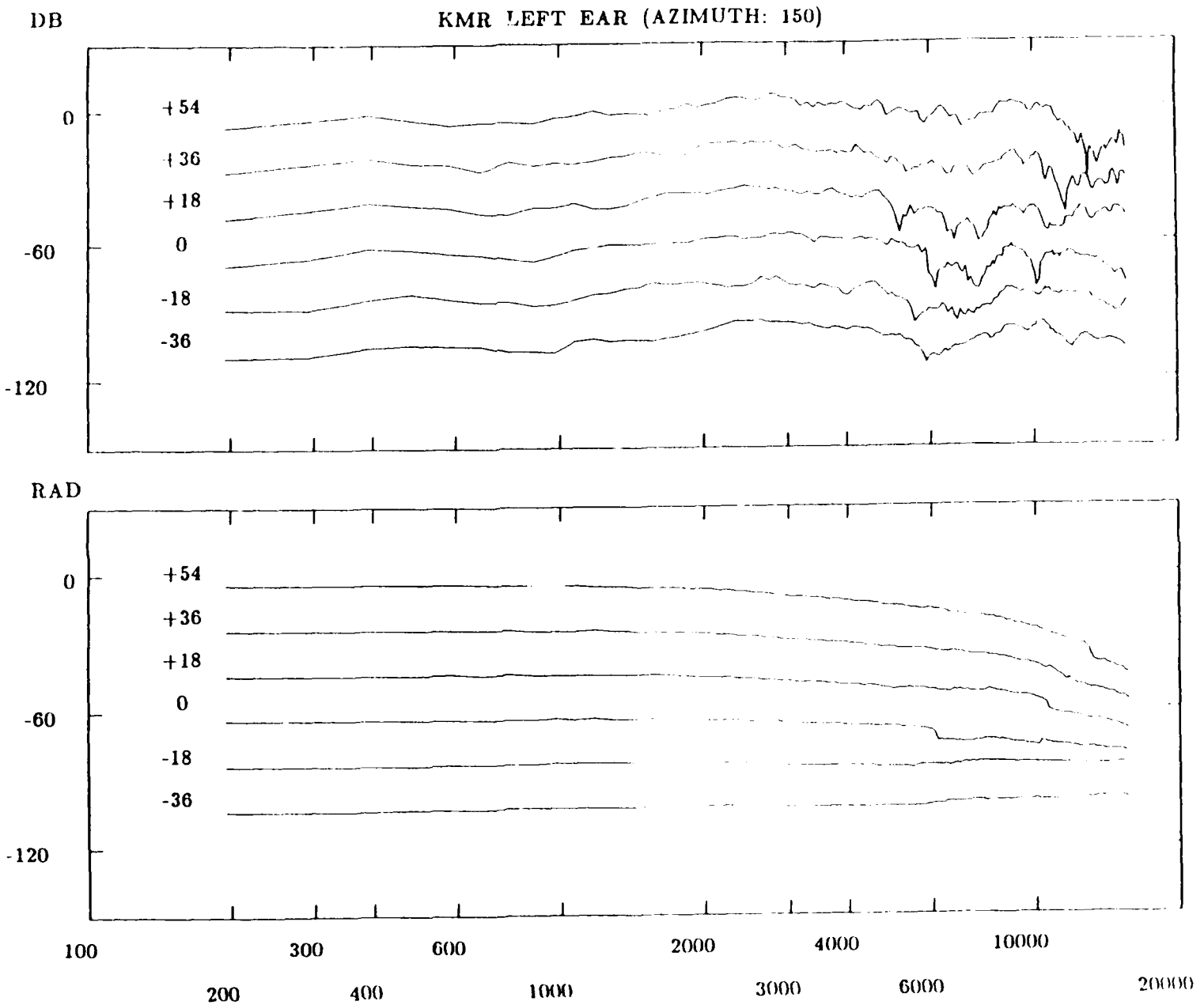


Figure 39. Same as Figure 29, except for a source at 150 degrees azimuth

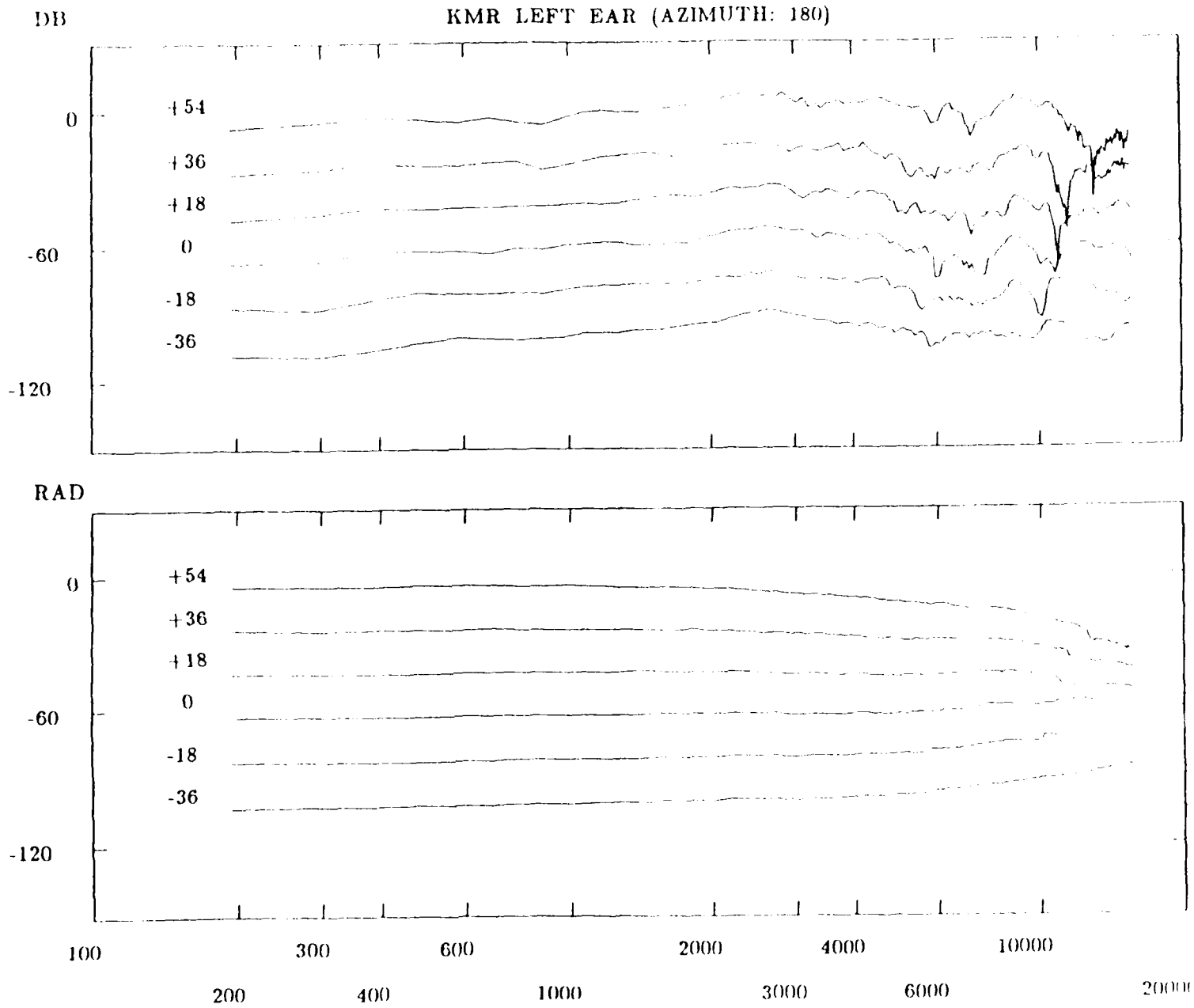


Figure 40. Same as Figure 29, except for a source at 180 degrees azimuth

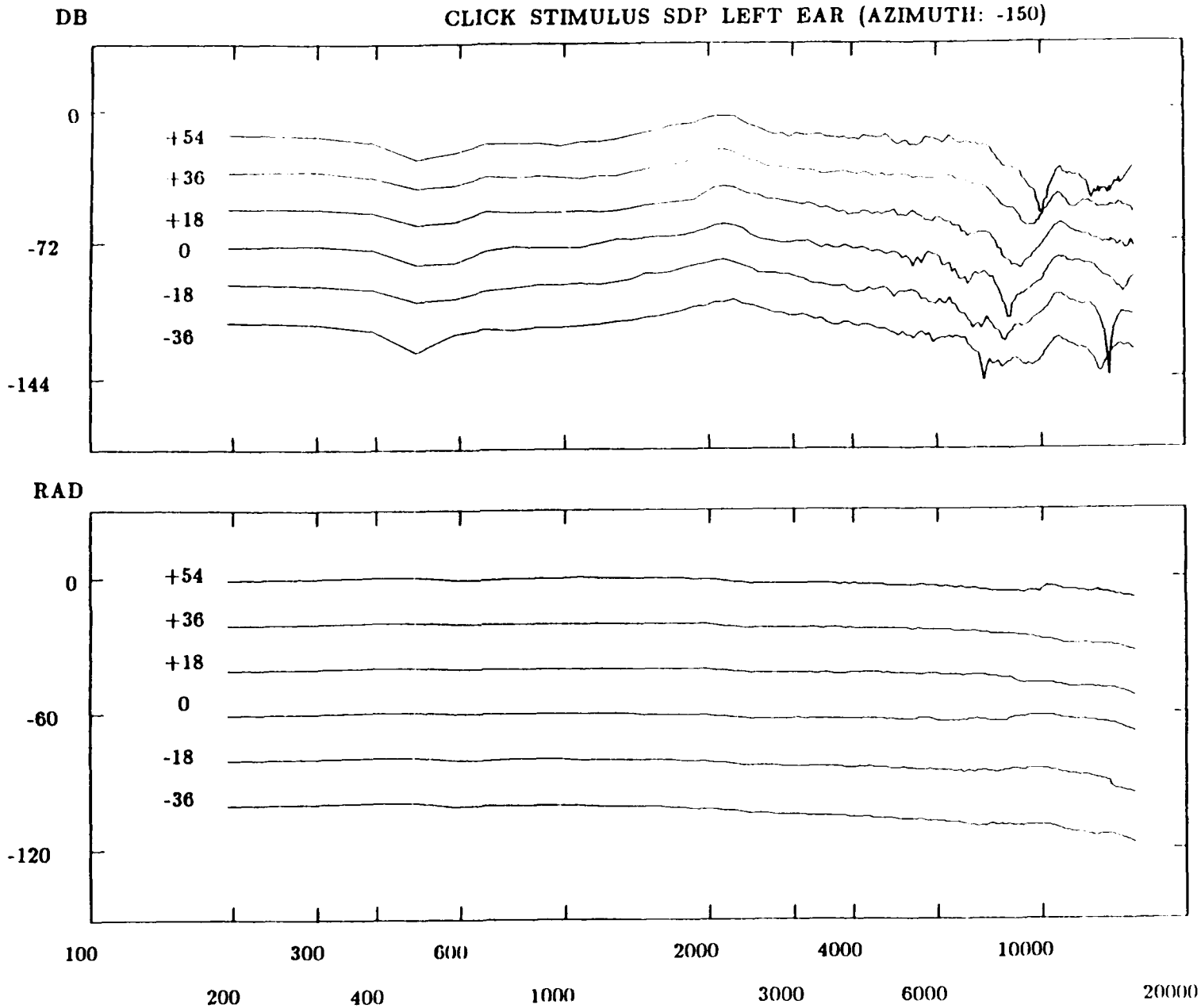
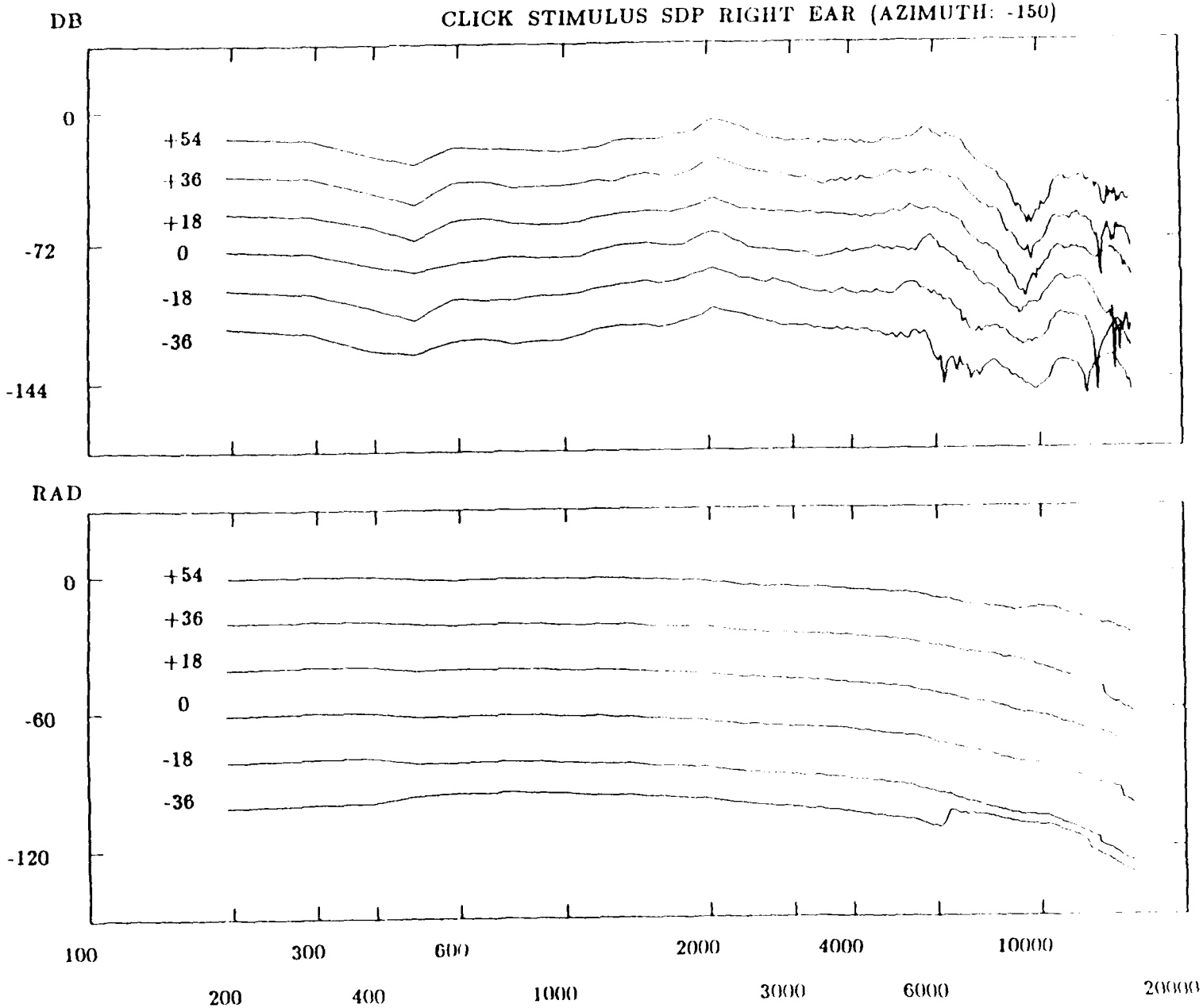


Figure 41. HRTF measurements from subject SDP for a source at -150 degrees azimuth (i.e., 150 degrees on the subject's left side) and+54, +36, +18, 0, -18, -36 degrees elevation. The left ear HRTF is plotted in panel (a) and the right ear, in panel (b). The upper panel in (a) and (b) is the magnitude function in dB coordinates and the lower panel is the "unwrapped" phasefunction in radian coordinates. The functions for the six elevations are displaced vertically by 20 dB for visibility. The measurements were obtained using a click


DB
CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: -150)
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000

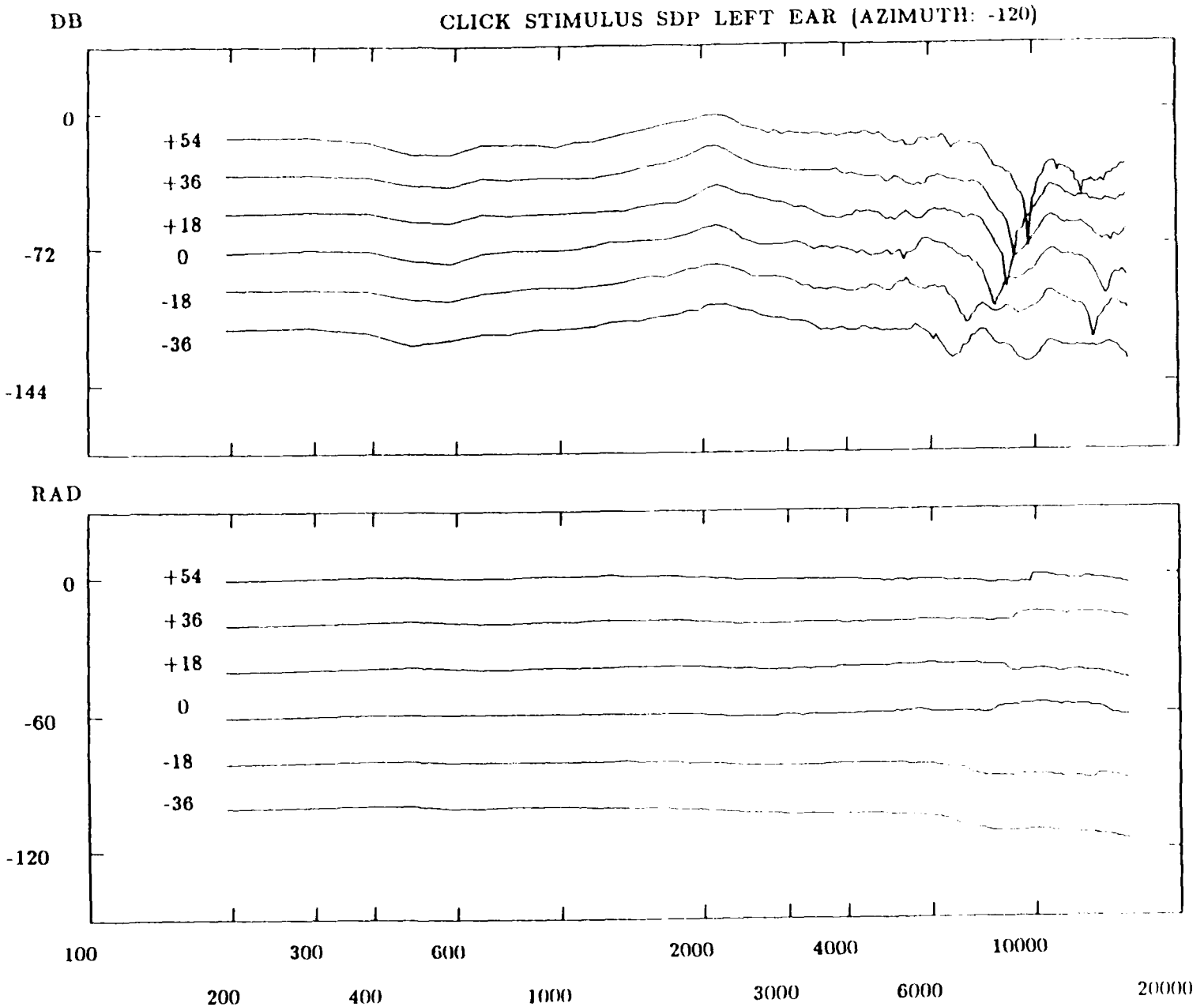
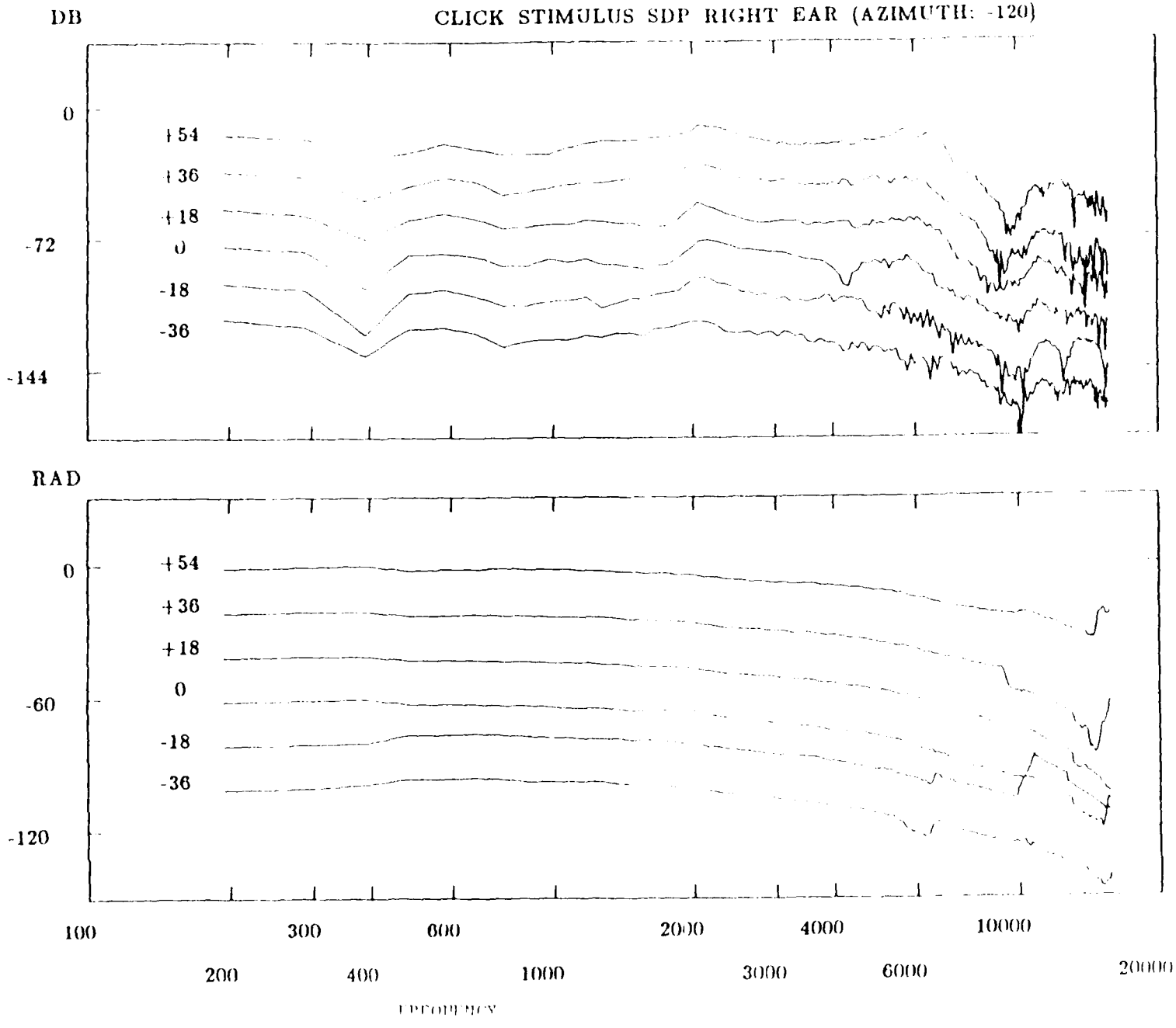


Figure 42. Same as Figure 41, except for a source at -120 degrees azimuth


DB
CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: -120)
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

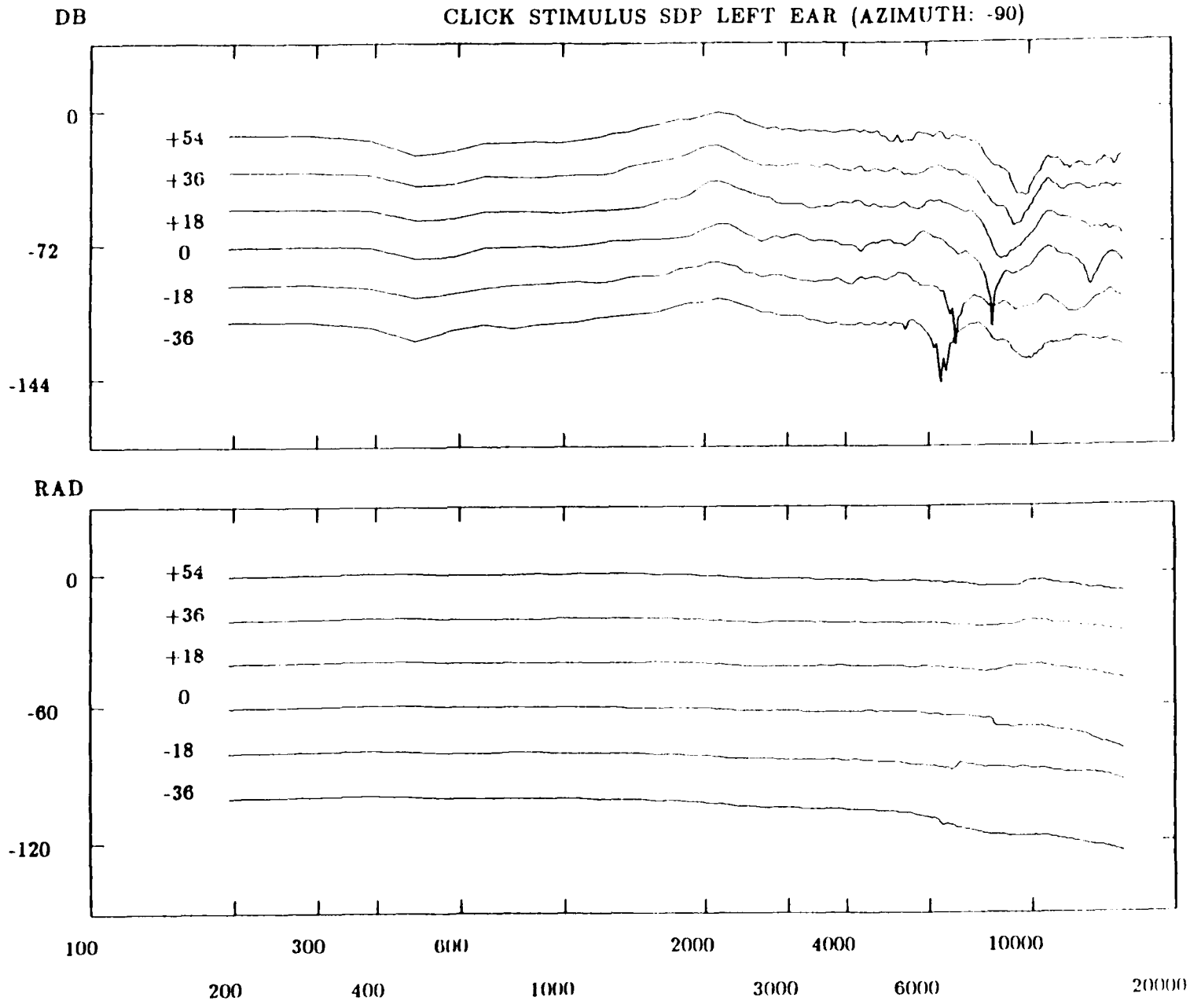
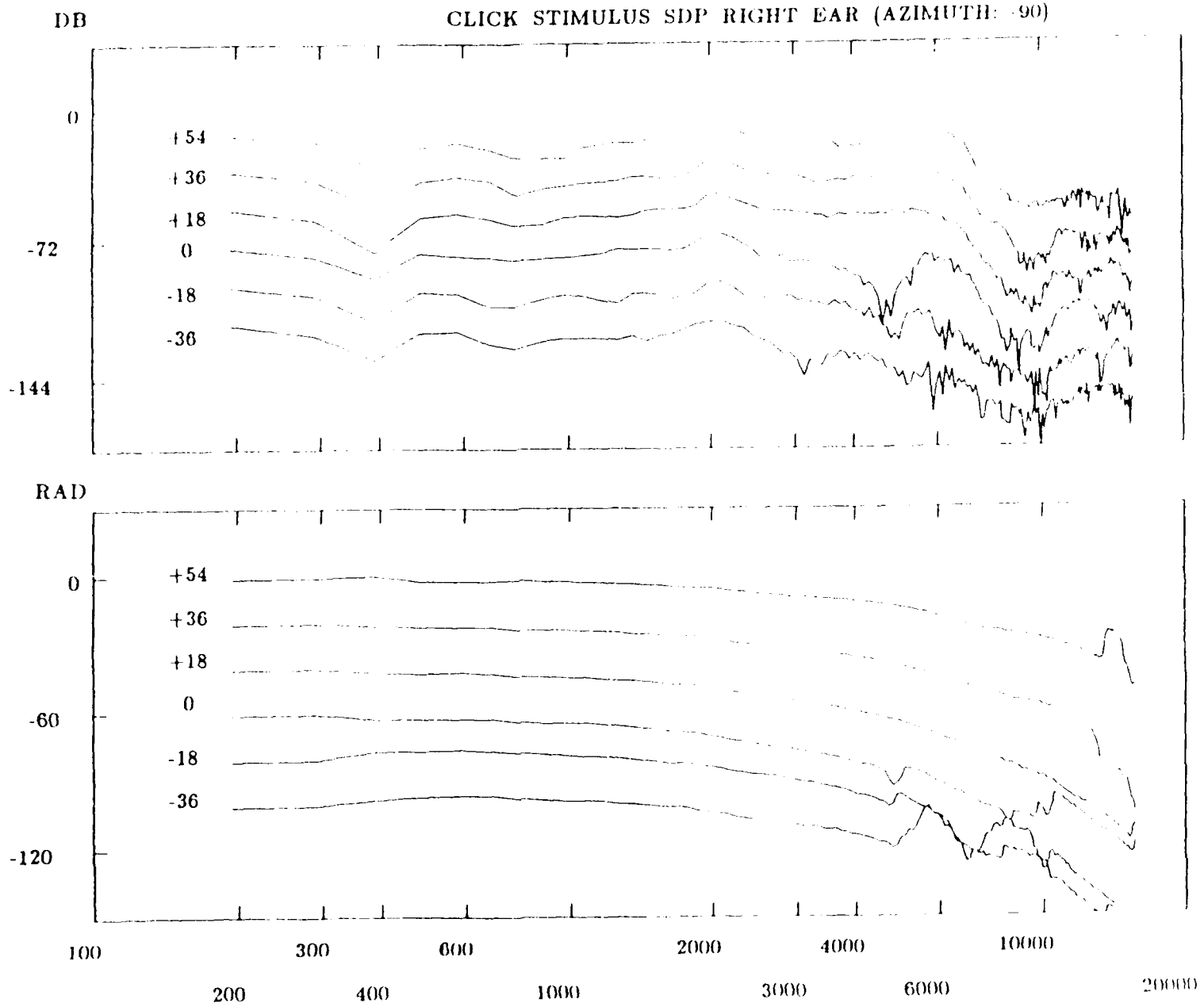


Figure 43. Same as Figure 41, except for a source at -90 degrees azimuth


DB
CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: -90)
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000

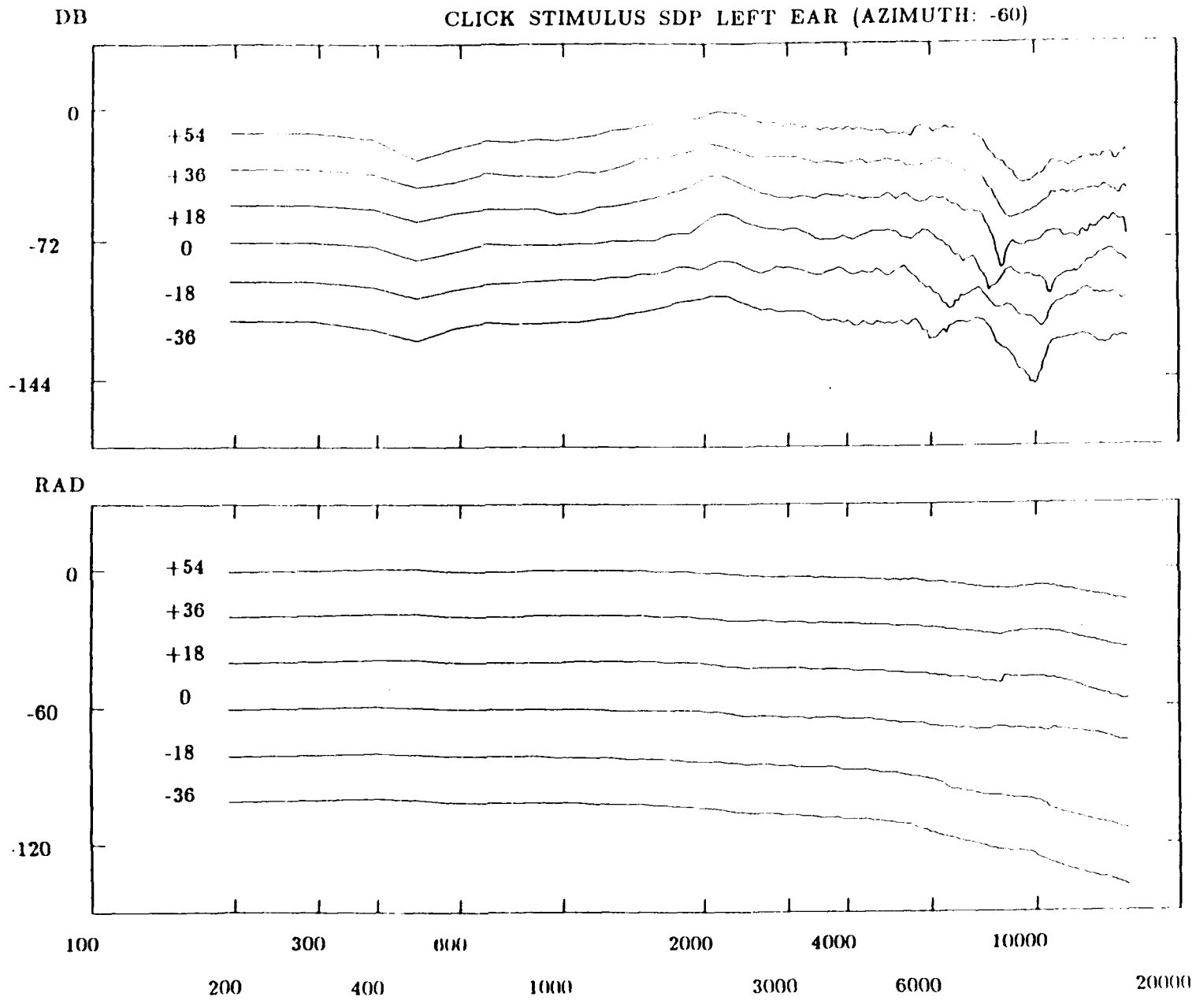
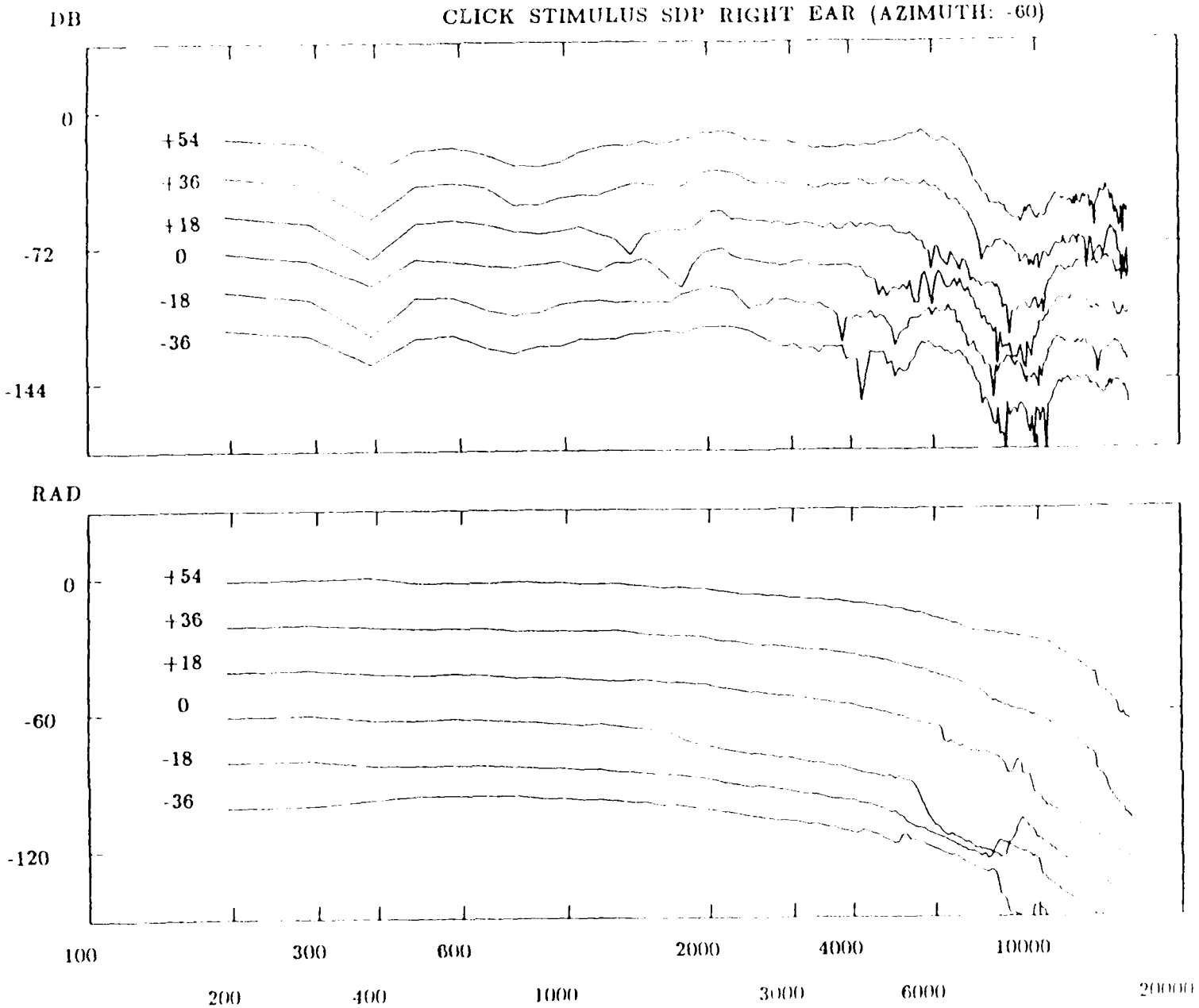


Figure 44. Same as Figure 41, except for a source at -60 degrees azimuth


DB
CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: -60)
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000

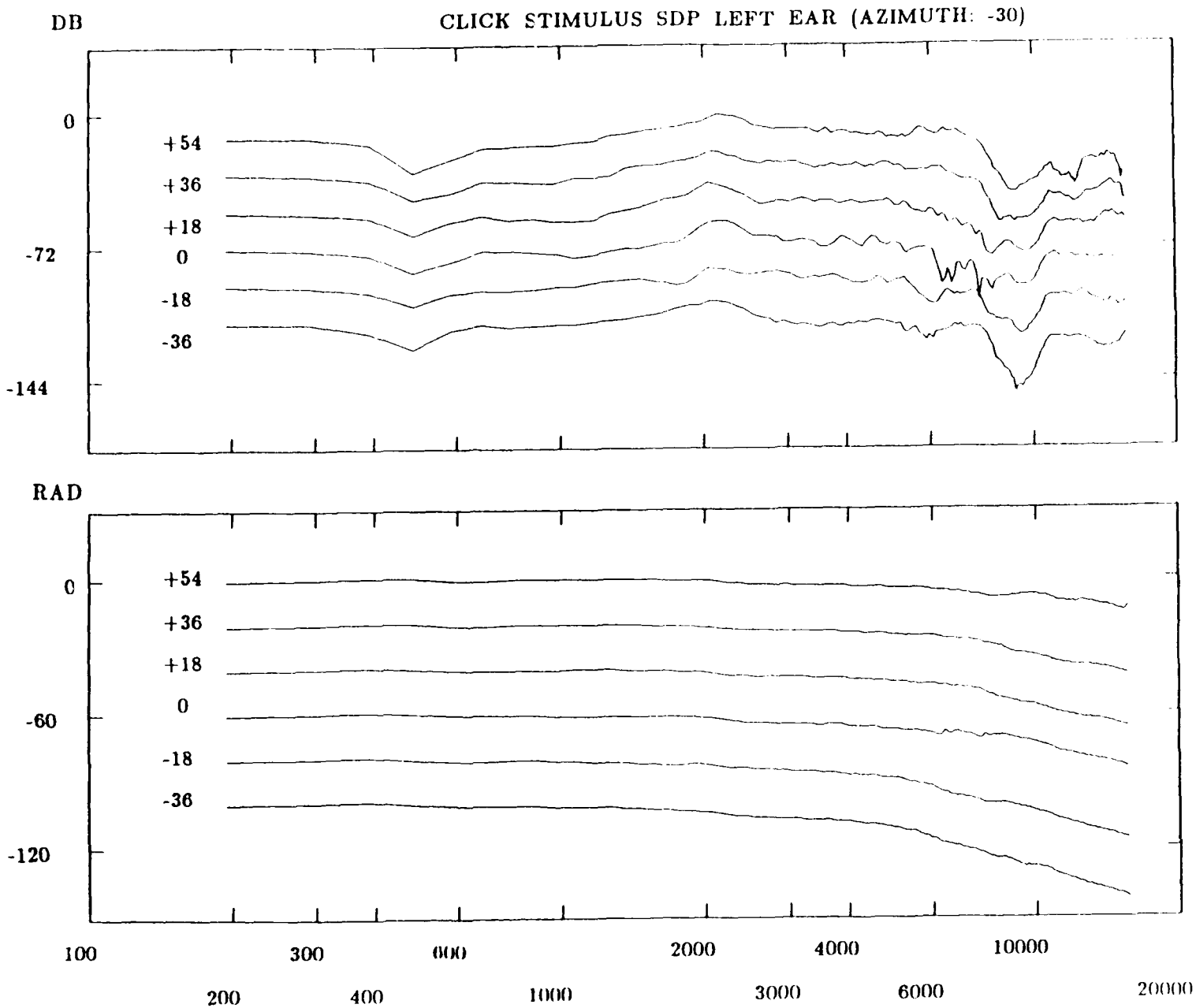
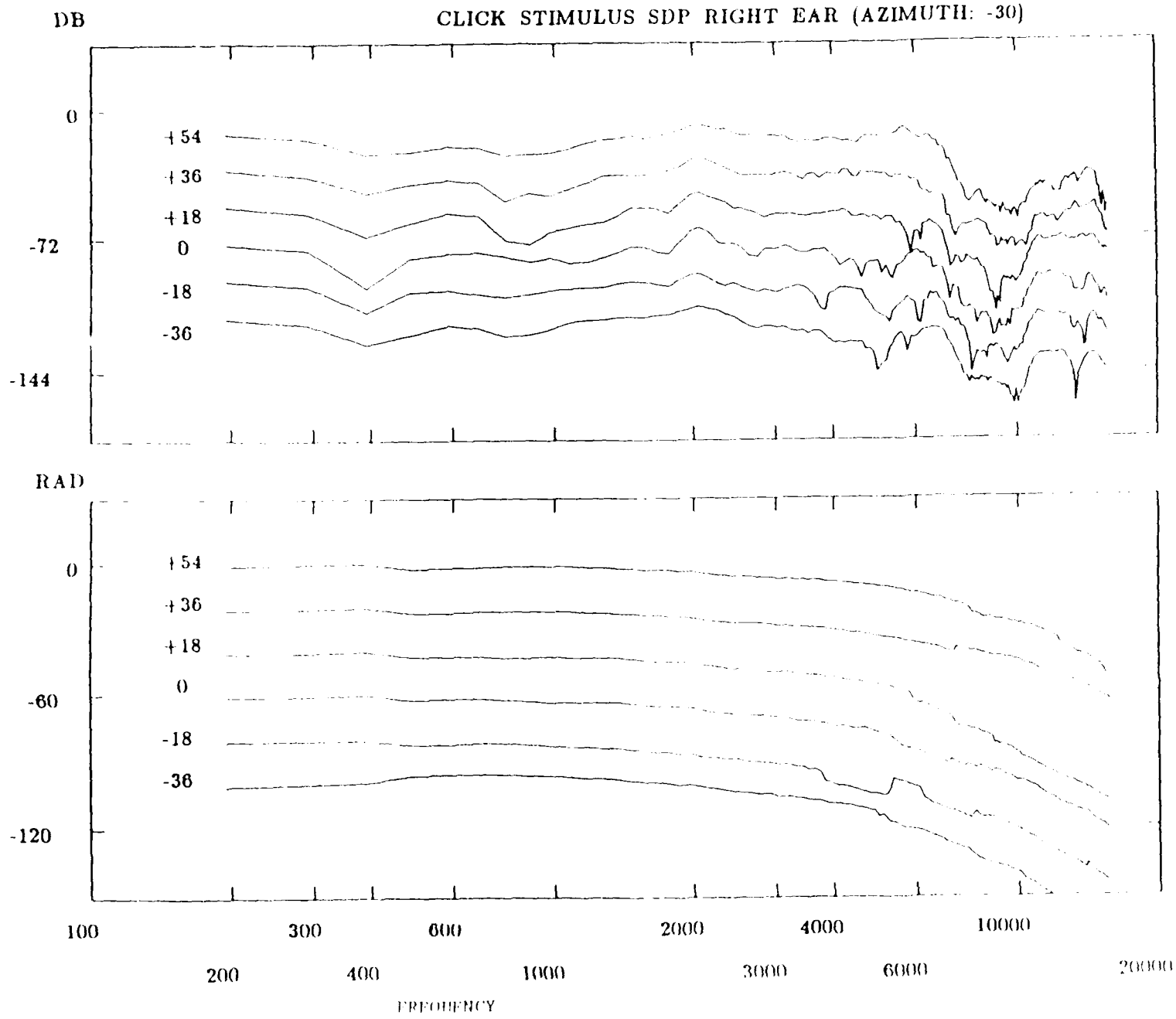


Figure 45. Same as Figure 41, except for a source at -30 degrees azimuth


DB
CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: -30)
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

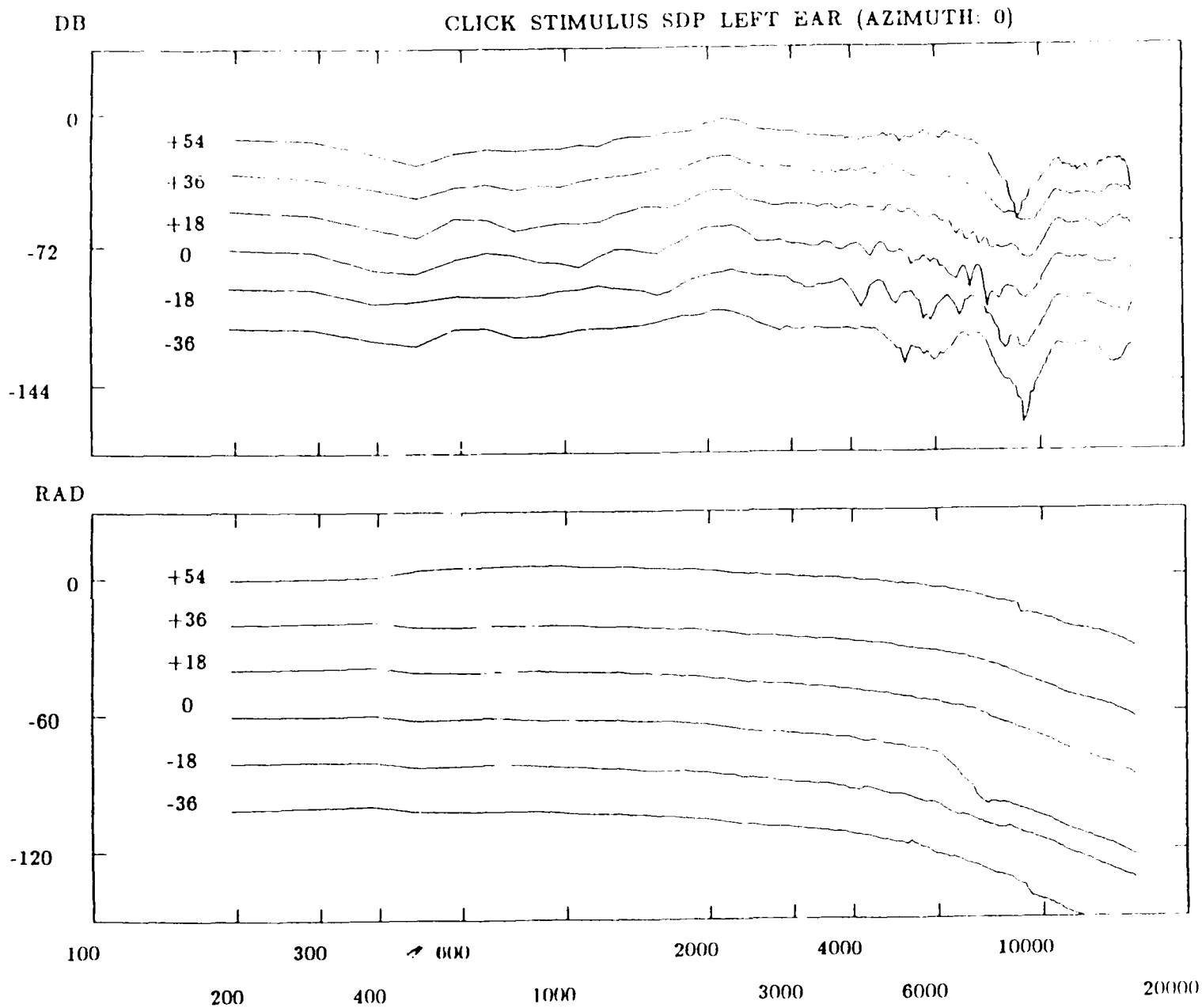
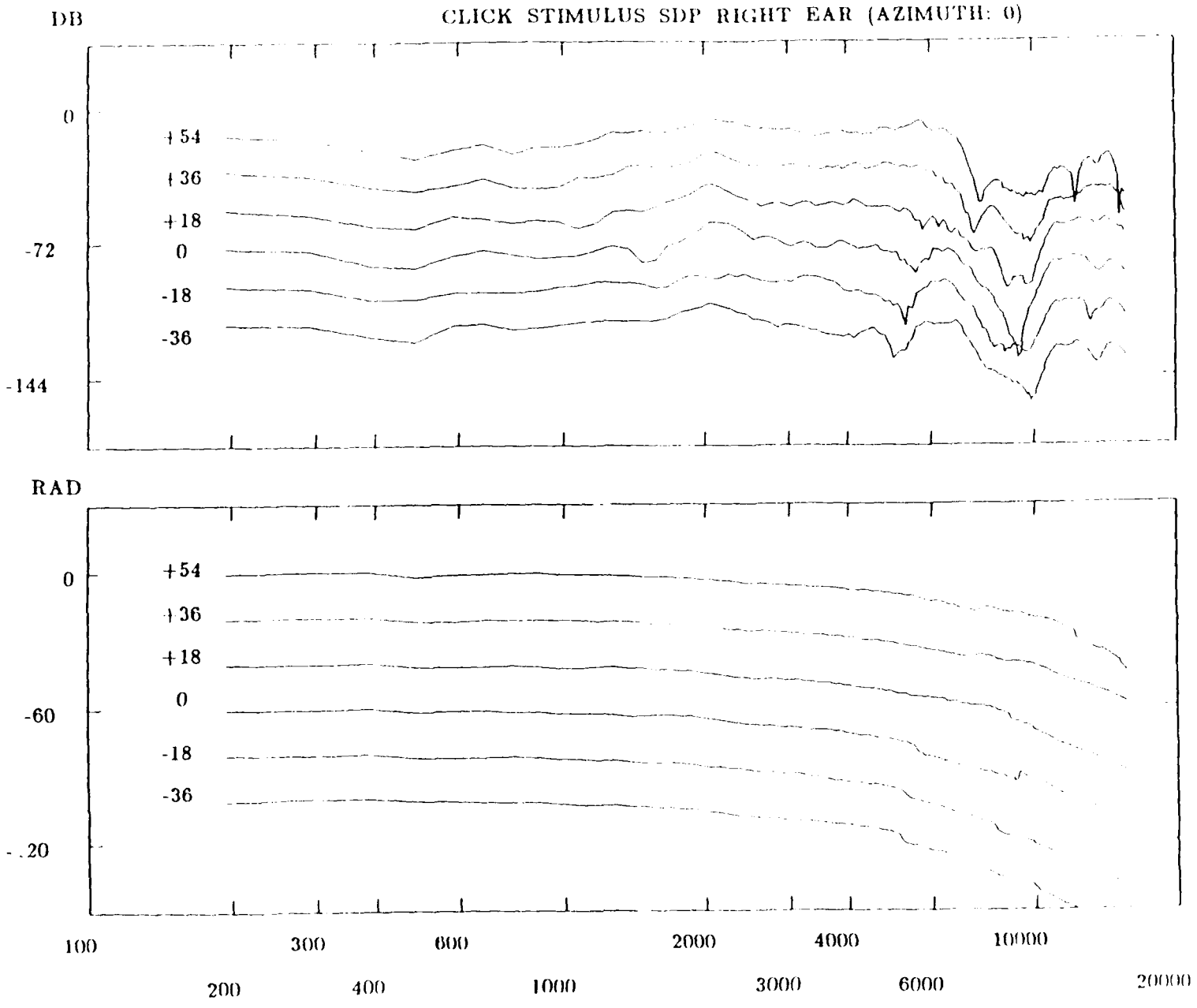


Figure 46. Same as Figure 41, except for a source at 0 degrees azimuth


DB
CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: 0)
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-.20
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000

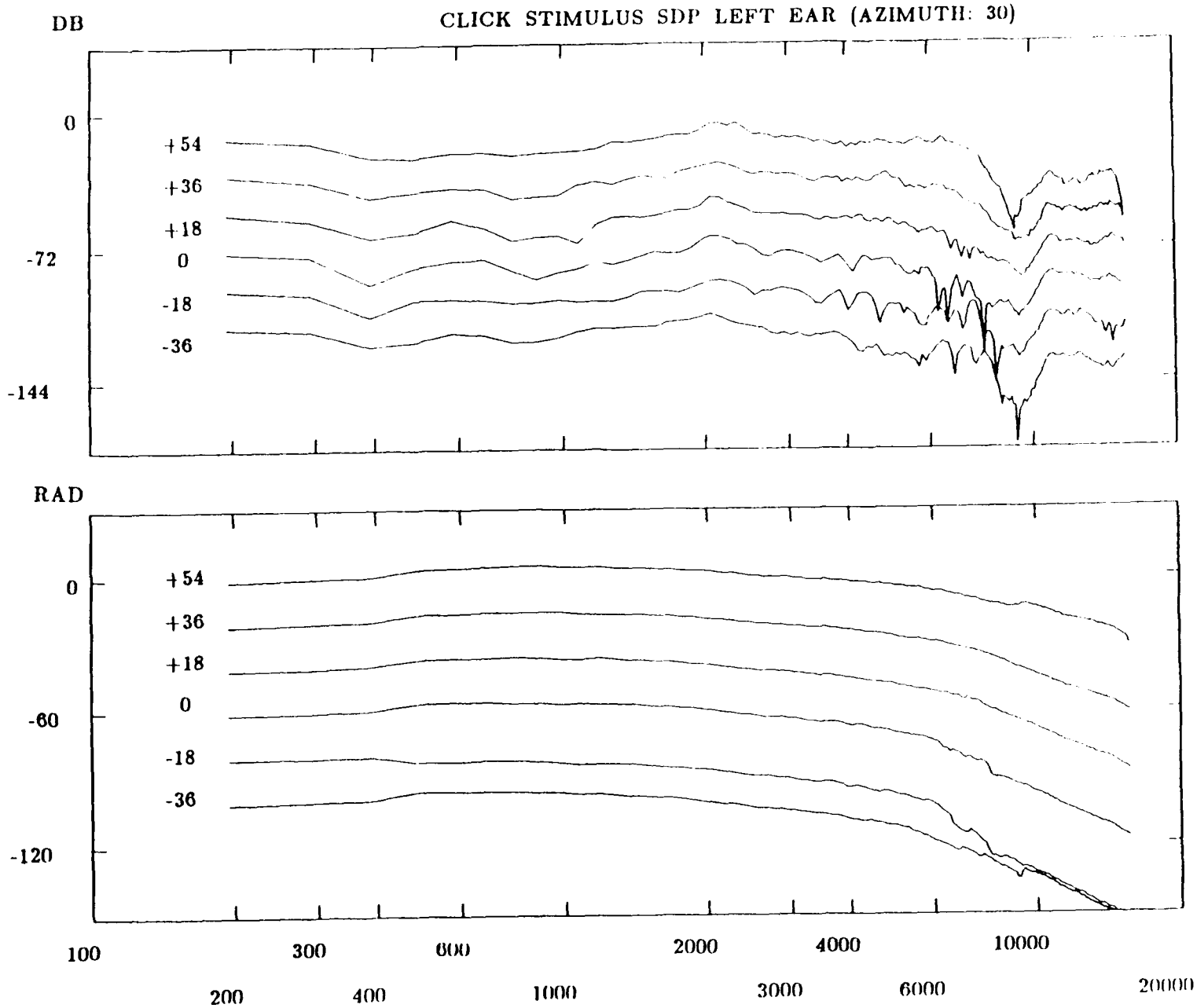
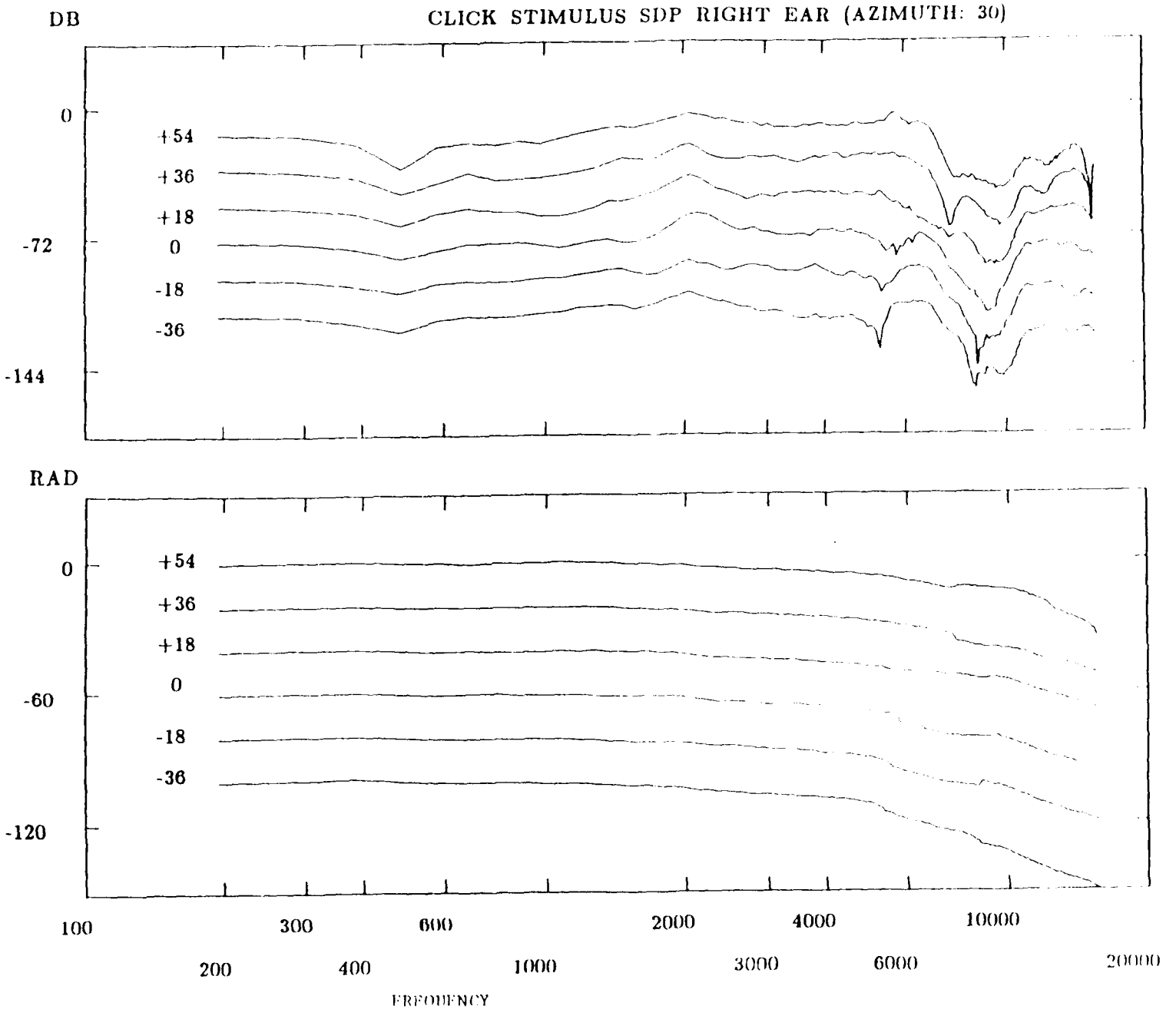


Figure 47. Same as Figure 41, except for a source at 30 degrees azimuth


DB
CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: 30)
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

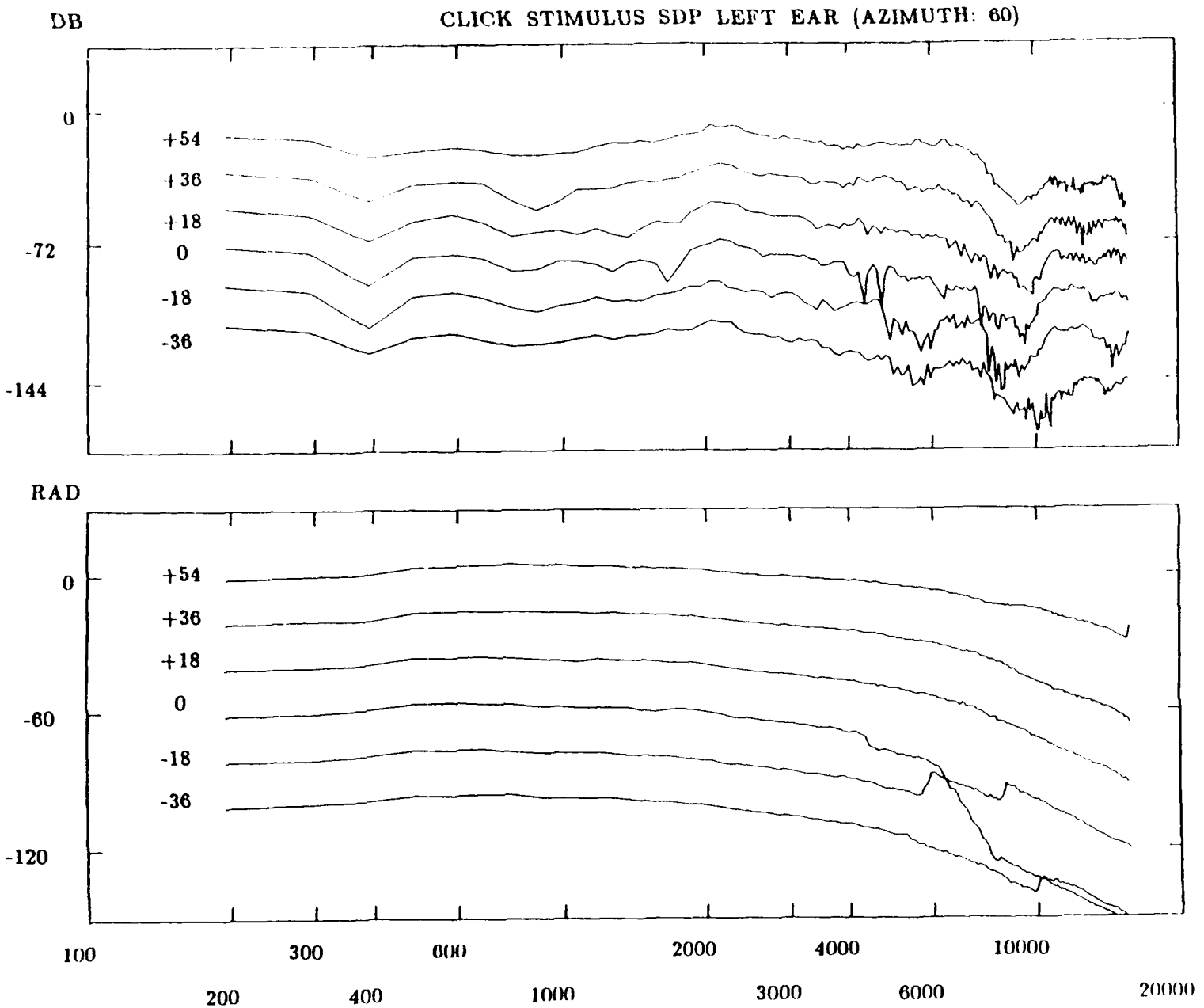
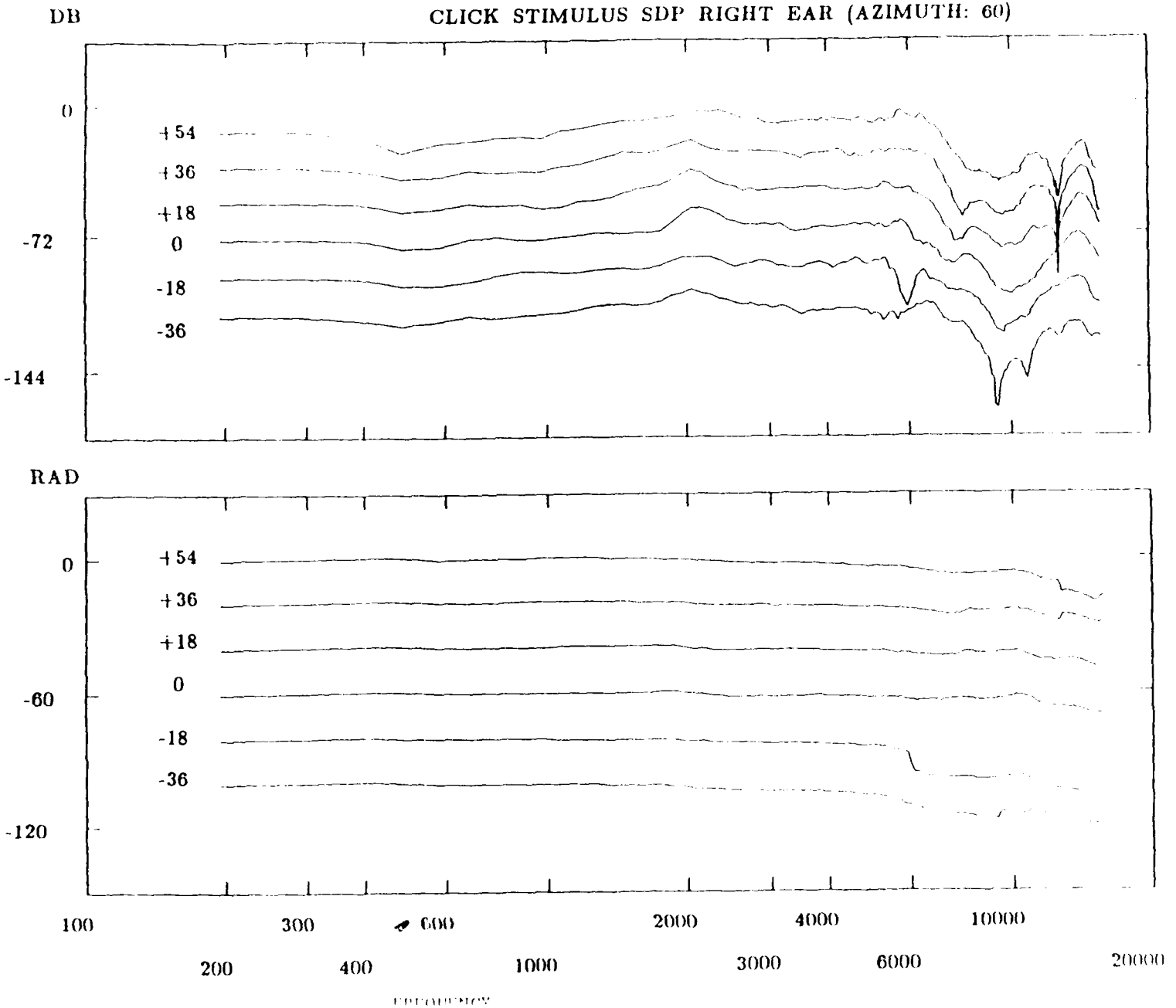


Figure 48. Same as Figure 41, except for a source at 60 degrees azimuth


DB
CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: 60)
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000

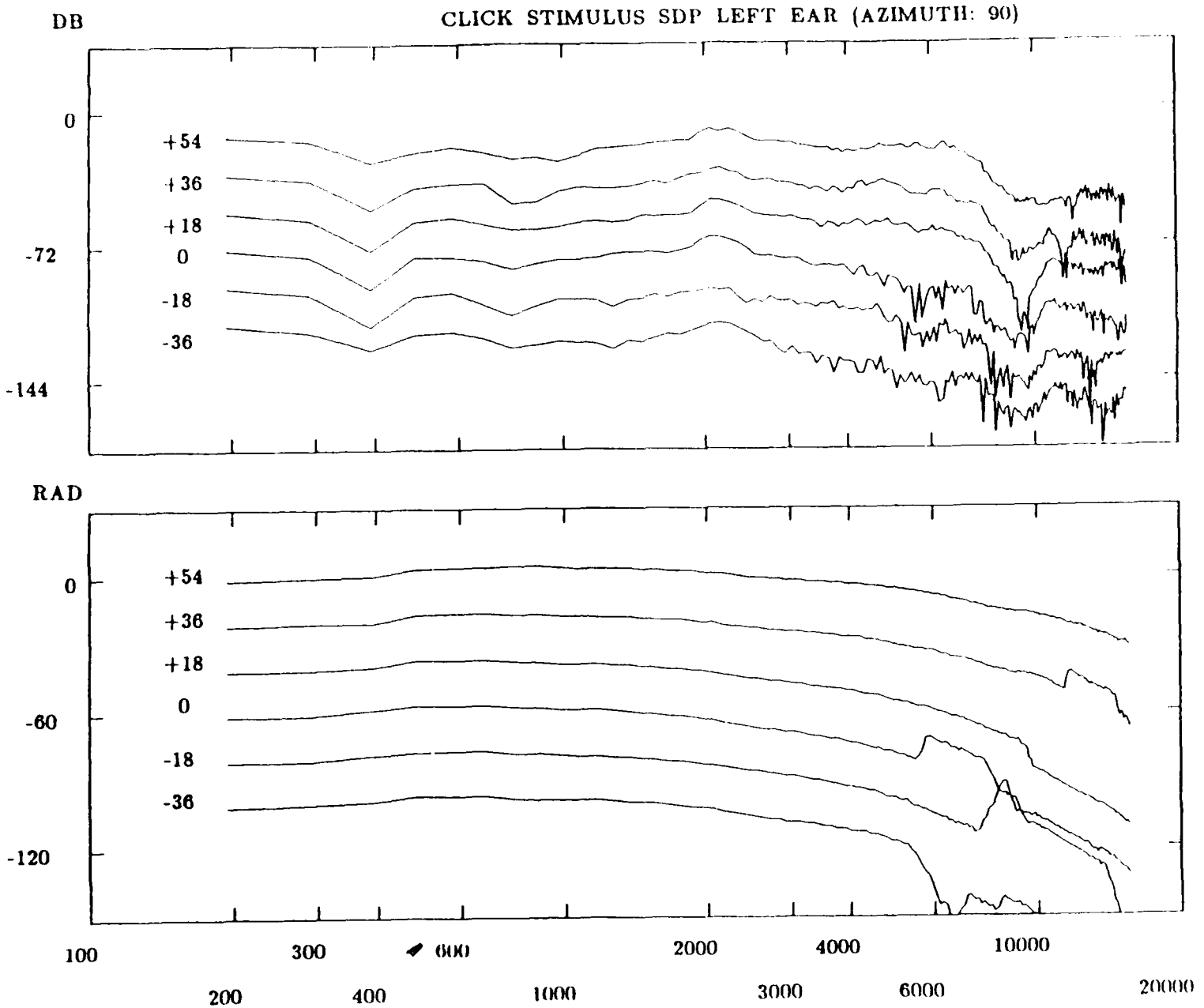
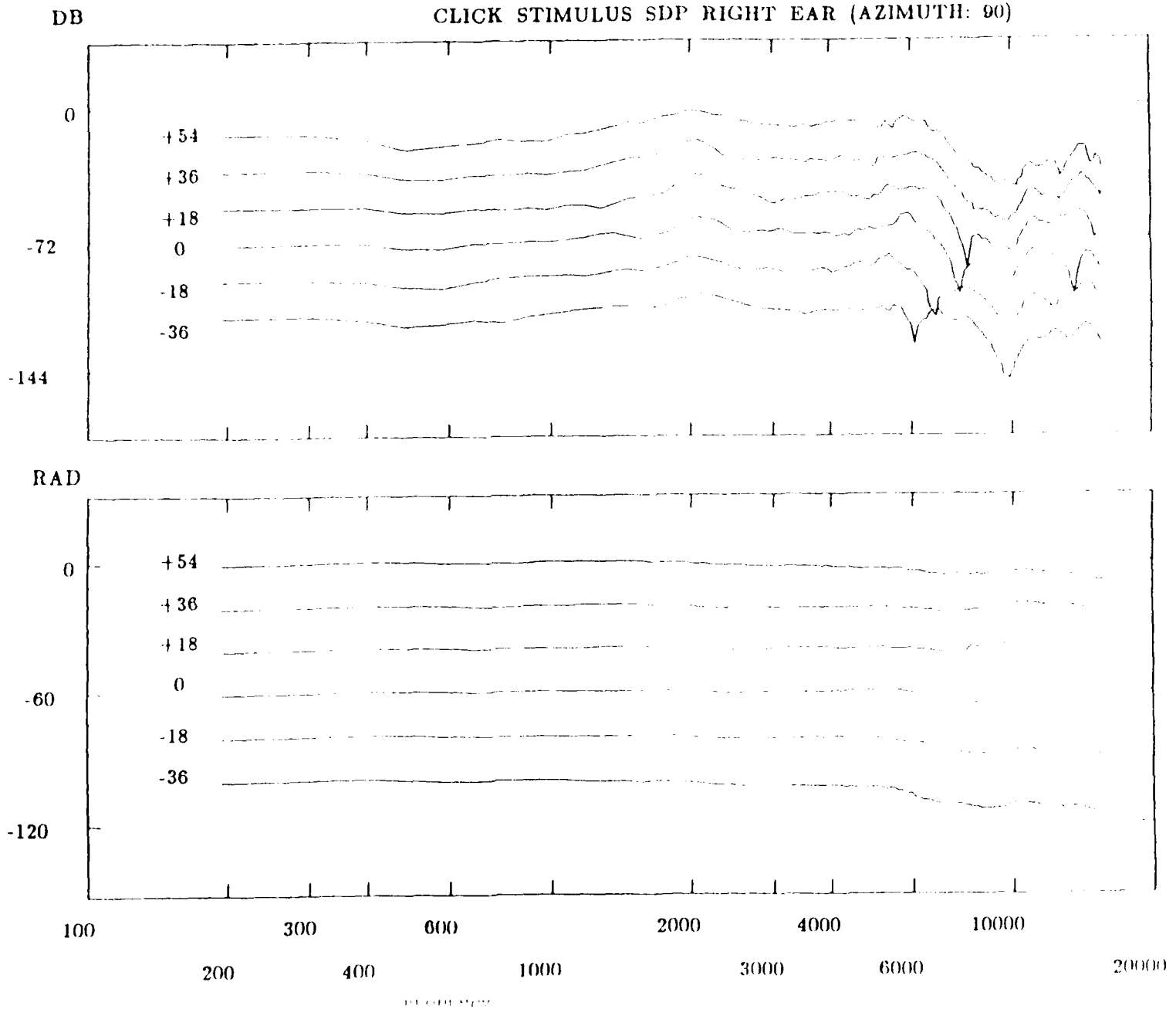


Figure 49. Same as Figure 41, except for a source at 90 degrees azimuth


DB
CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: 90)
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000

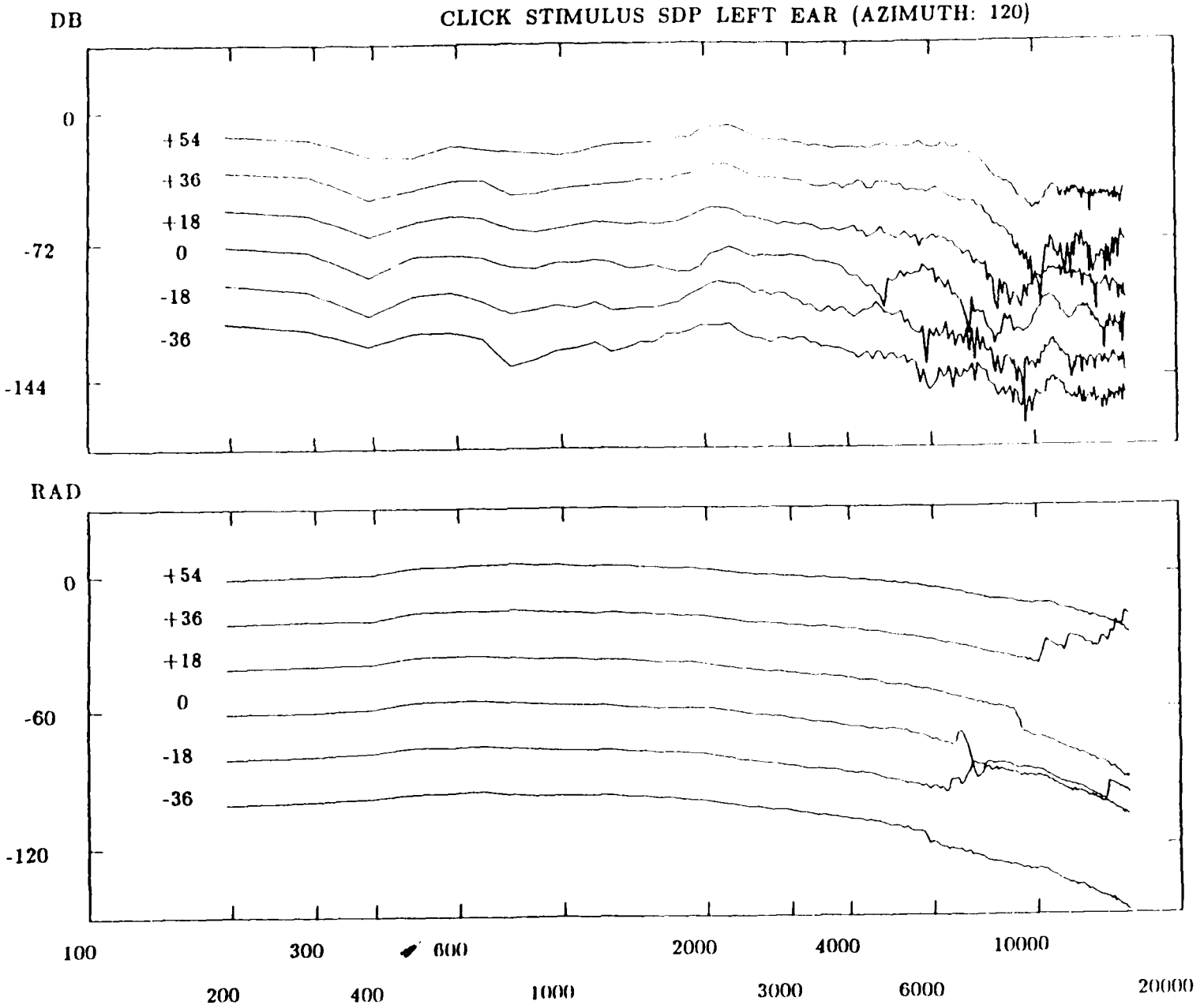
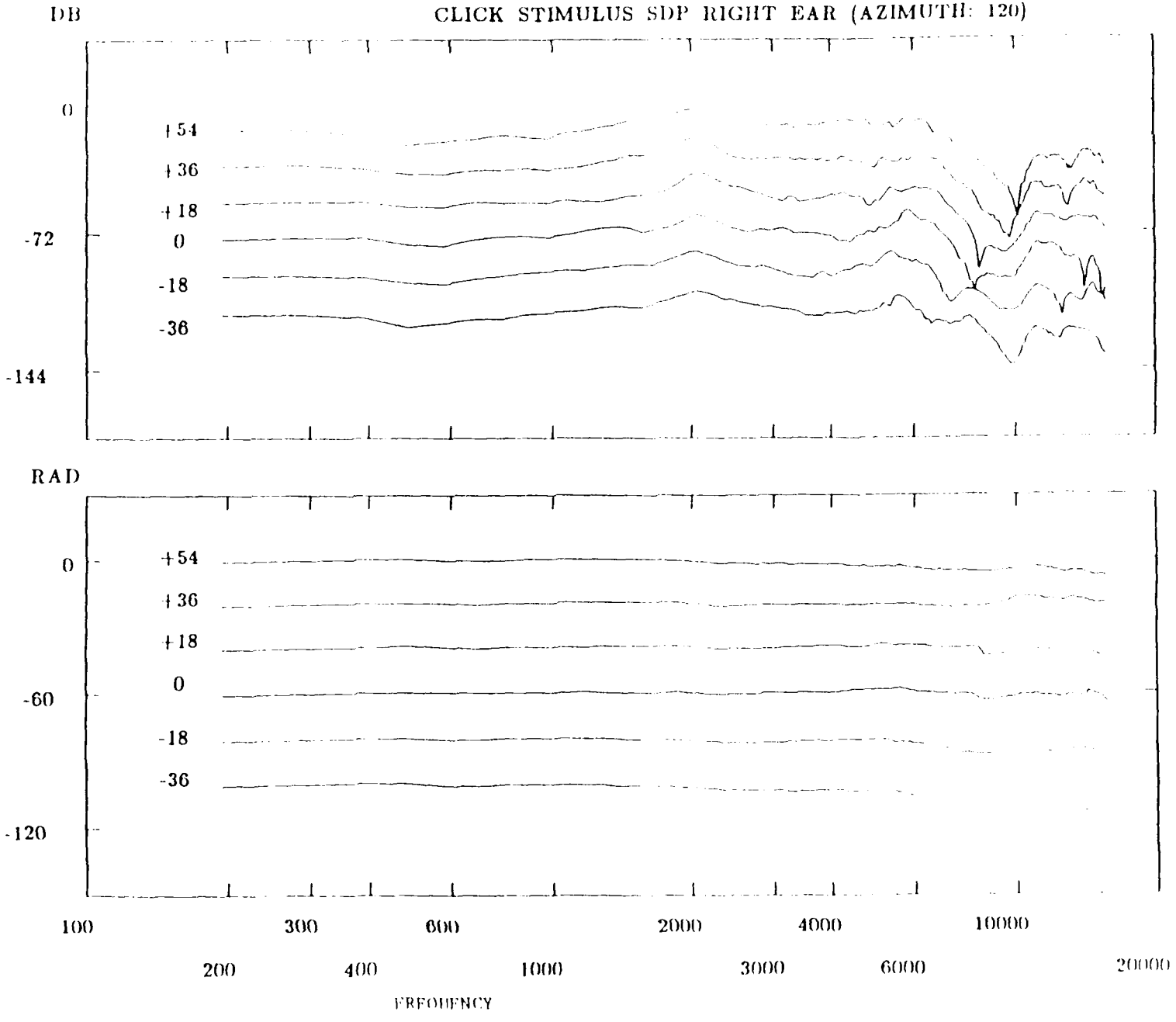


Figure 50. Same as Figure 41, except for a source at 120 degrees azimuth


DB
CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: 120)
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

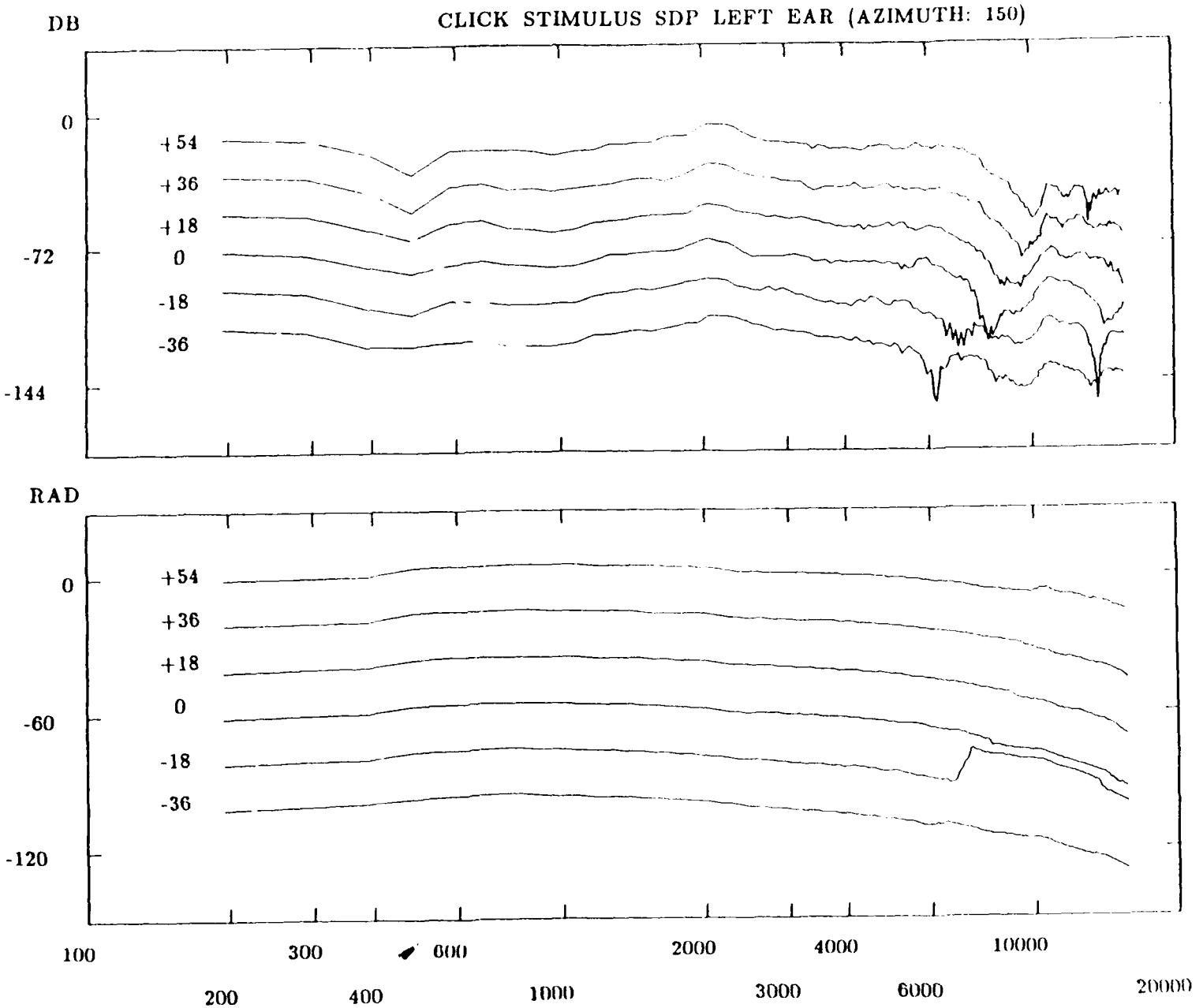
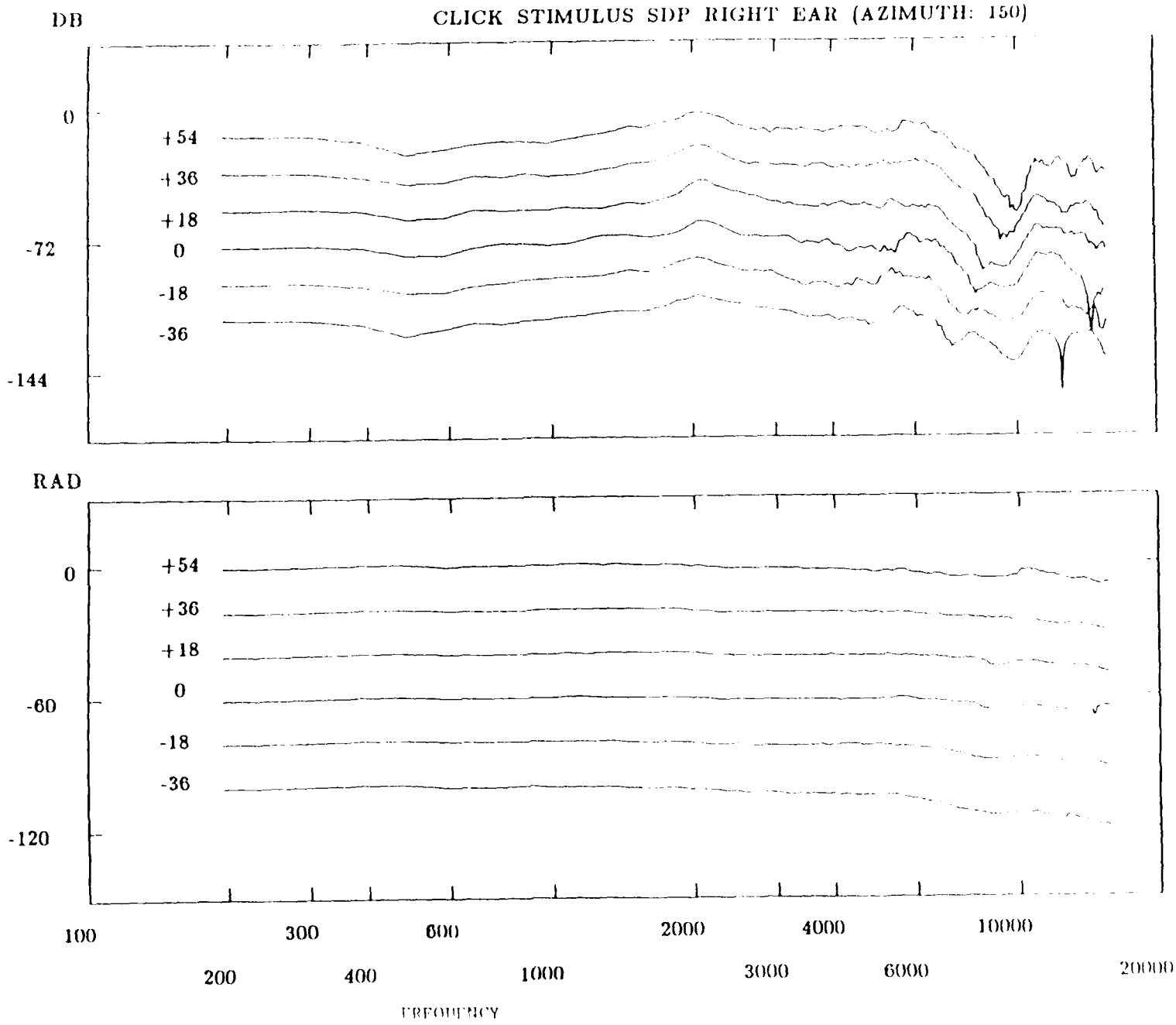


Figure 51. Same as Figure 41, except for a source at 150 degrees azimuth


CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: 150)
DB
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

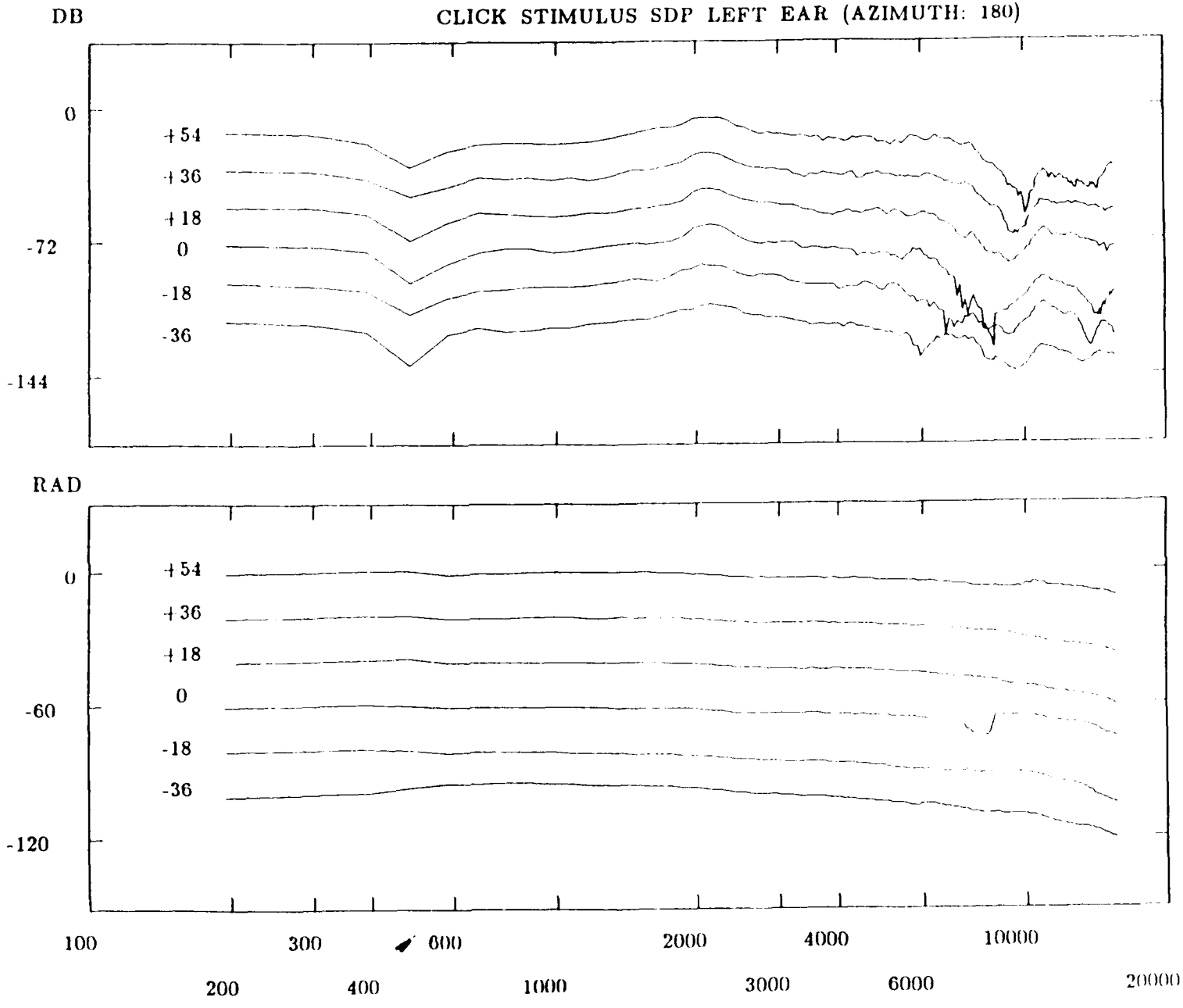
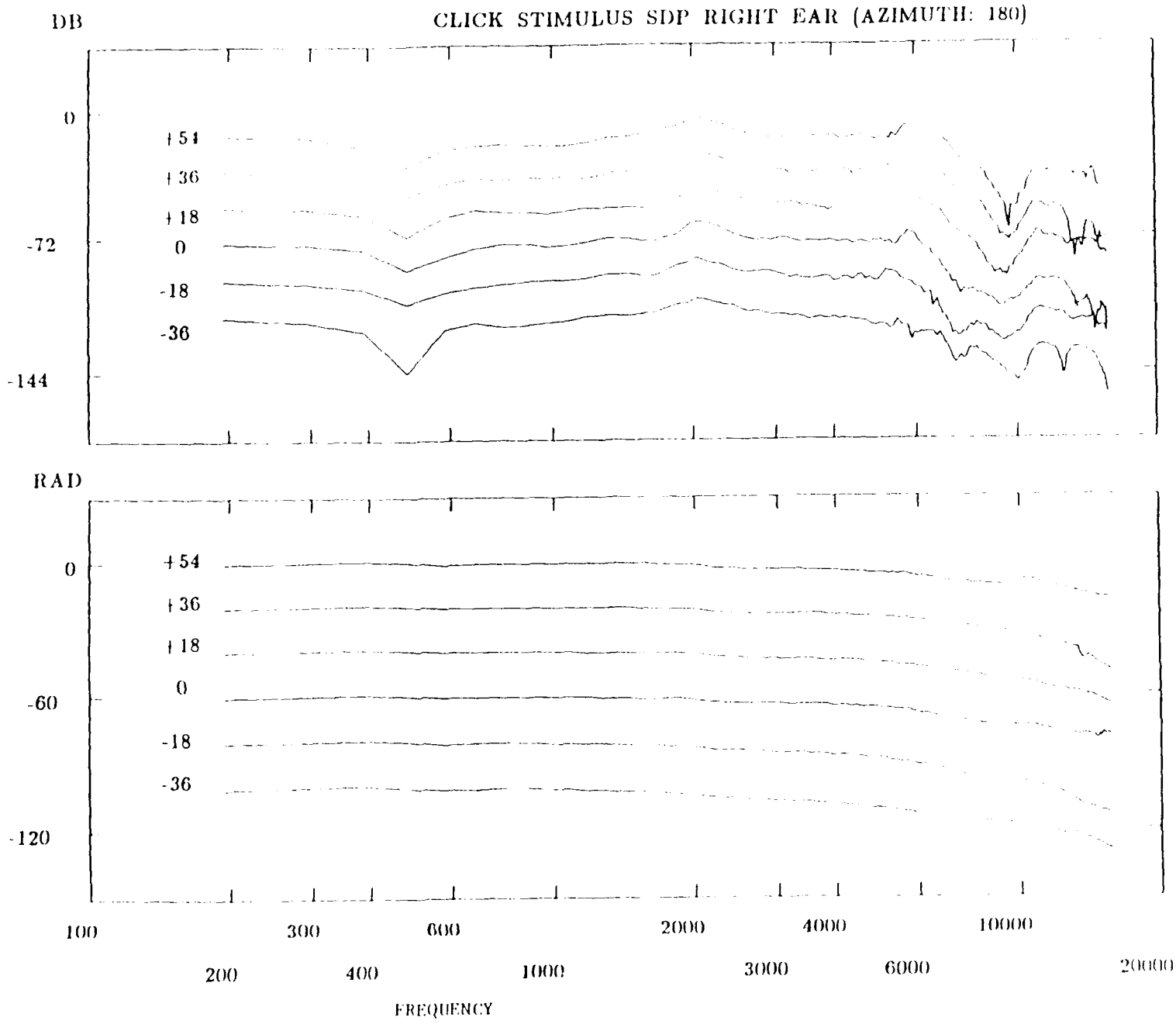


Figure 52. Same as Figure 41, except for a source at 180 degrees azimuth


CLICK STIMULUS SDP RIGHT EAR (AZIMUTH: 180)
DB
0
-72
-144
+54
+36
+18
0
-18
-36
RAD
0
-60
-120
+54
+36
+18
0
-18
-36
100
200
300
400
600
1000
2000
3000
4000
6000
10000
20000
FREQUENCY

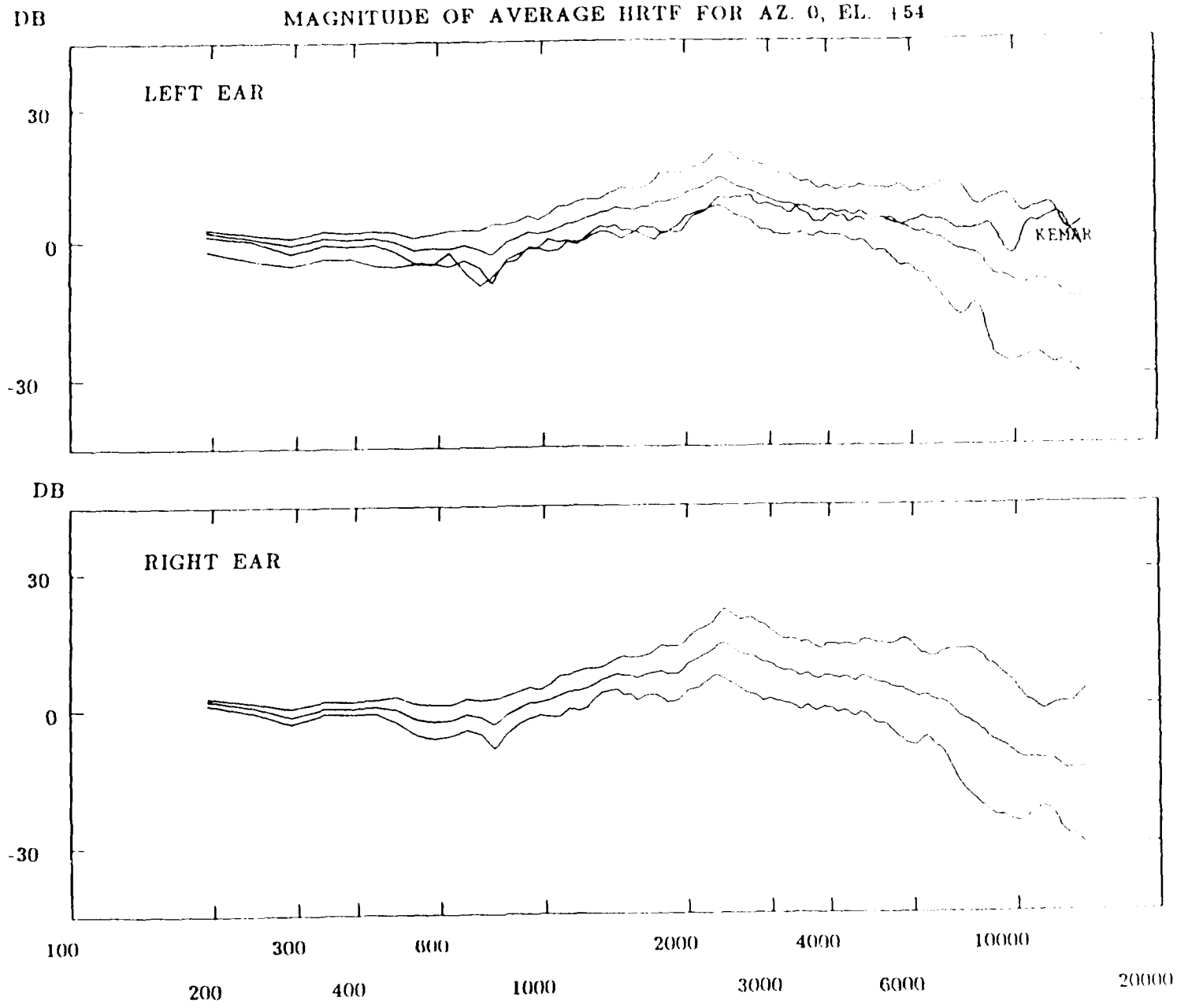


Figure 53. Average magnitude function and 95% confidence limits of HRTFs of 20 Subjects at 0 degrees azimuth and +54 degrees elevation. The magnitude of the HRTF of KEMAR's left ear is also plotted

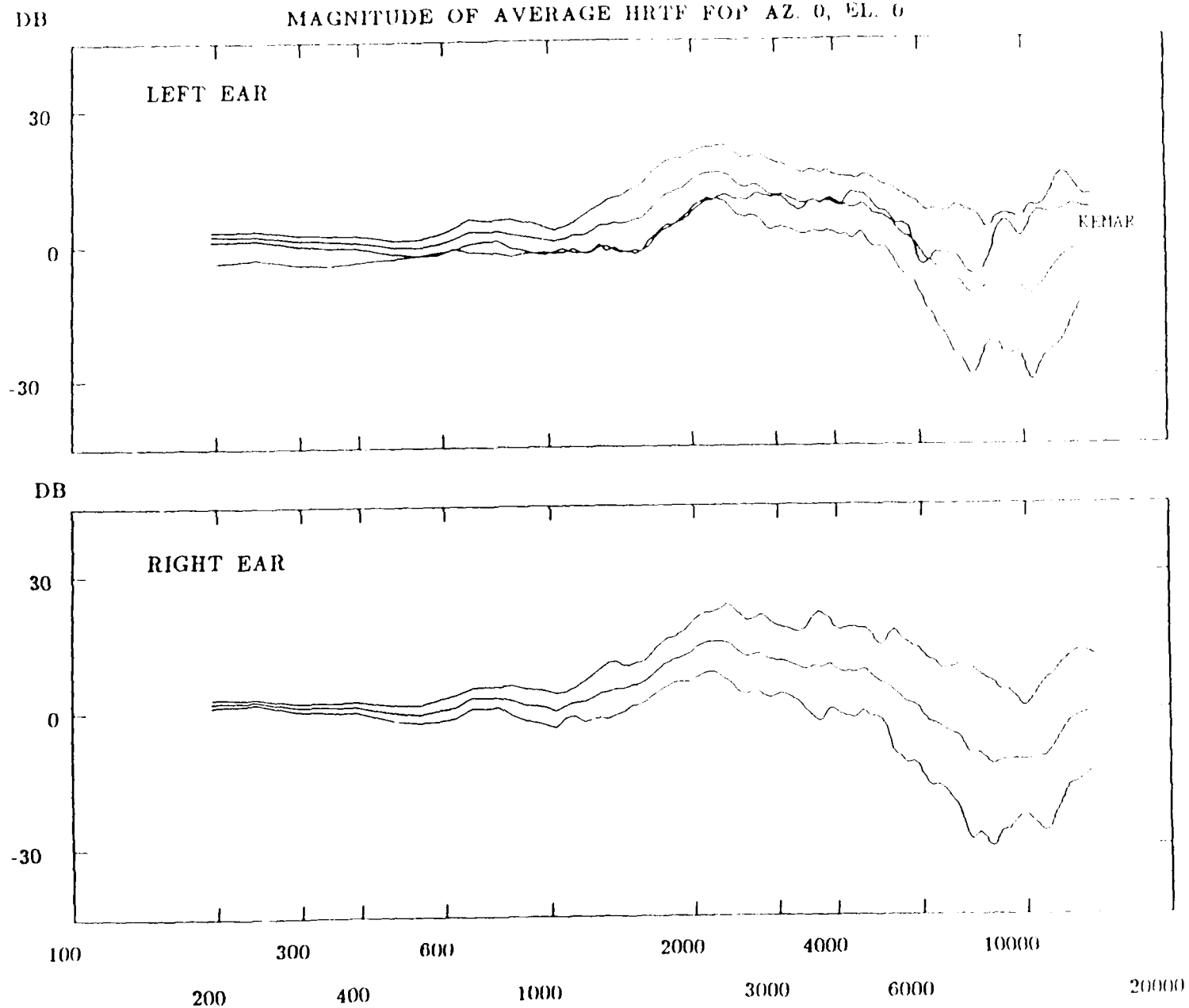


Figure 54. Same as Figure 53, except for 0 degrees azimuth and 0 degrees elevation

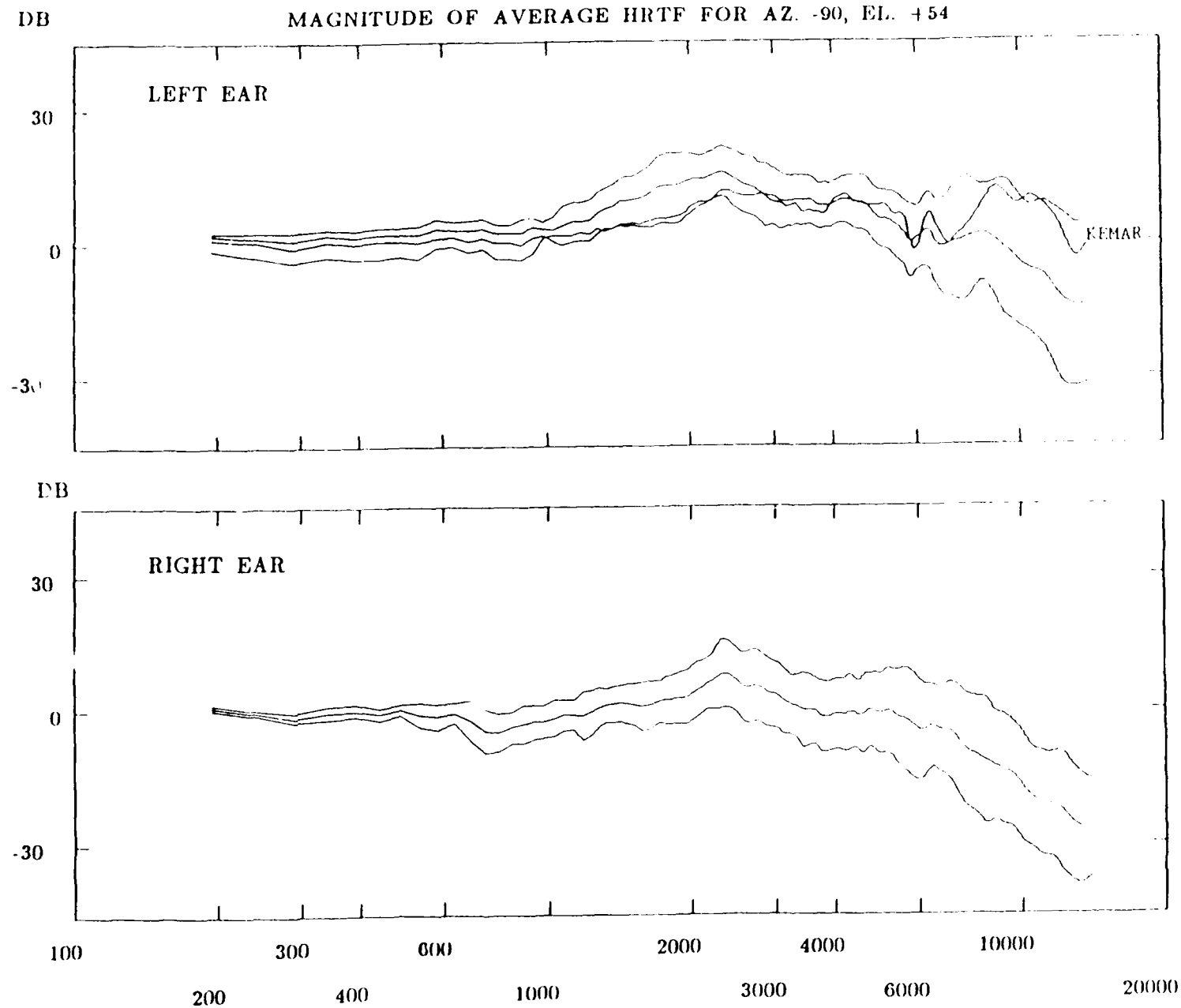


Figure 55. Same as Figure 53, except for -90 degrees azimuth and +54 degrees elevation

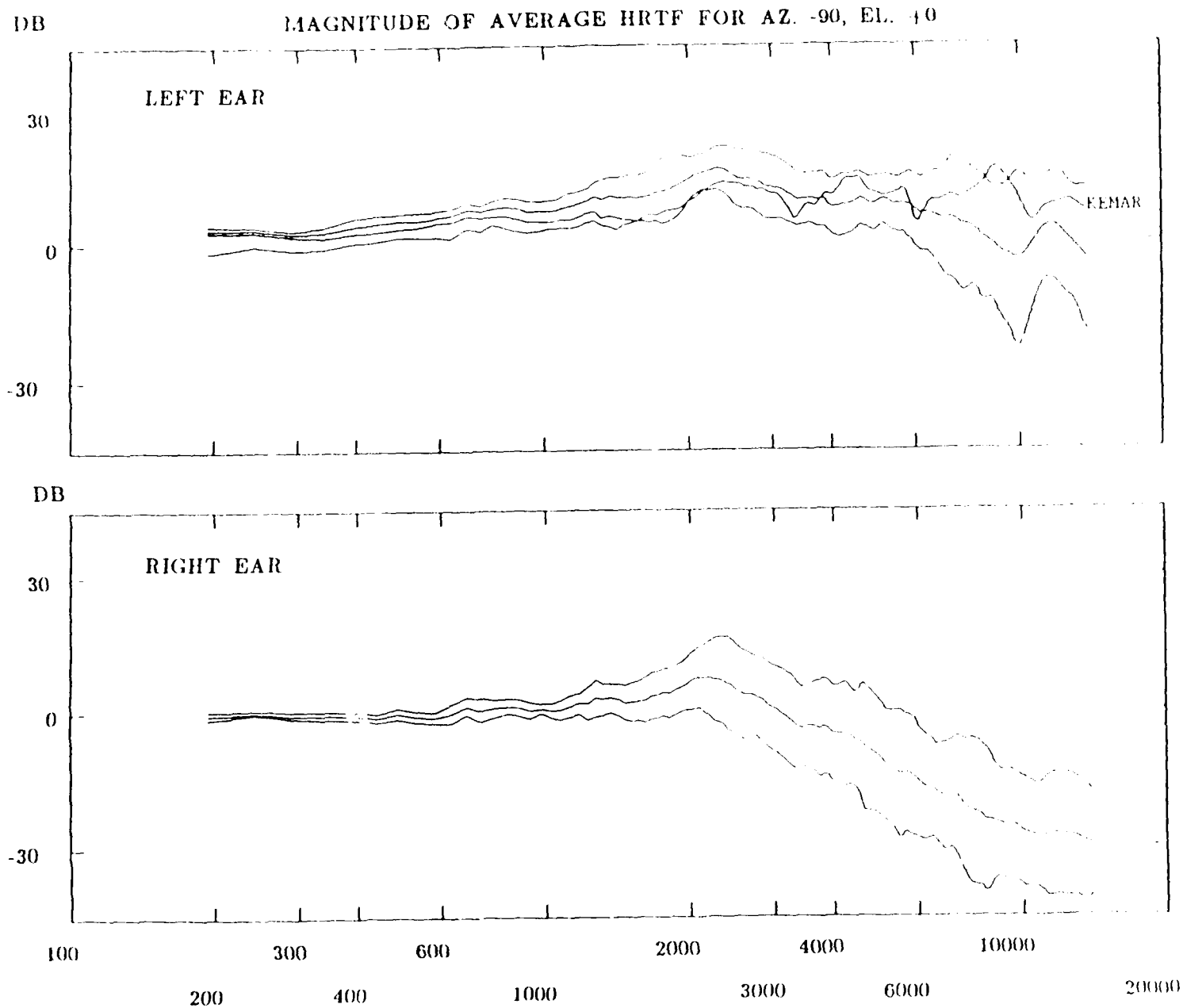


Figure 56. Same as Figure 53, except for -90 degrees azimuth and +54 elevation

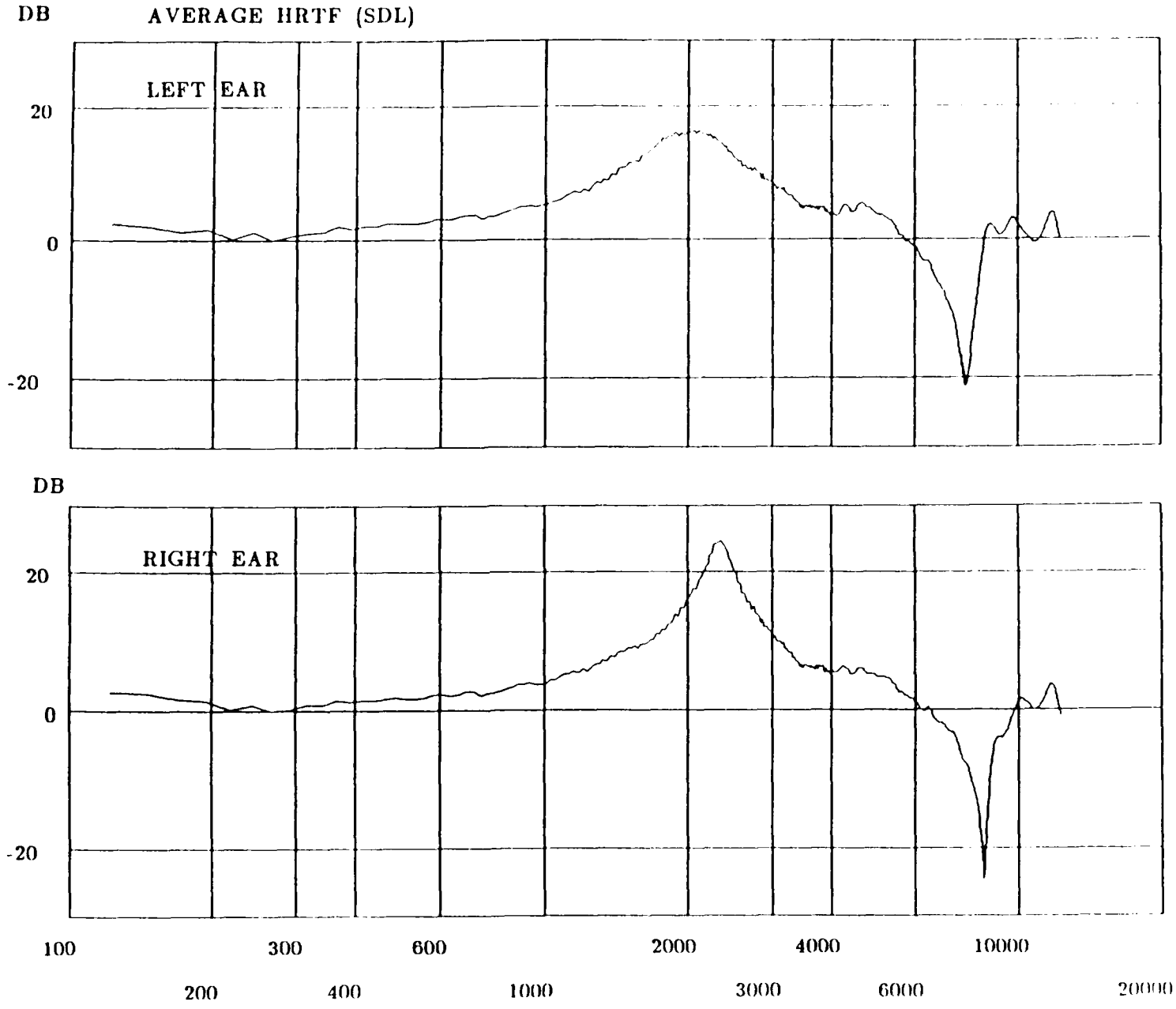


Figure 57. Mean of the magnitude of 144 HRTFs of Subject SDL. This average is an estimate of the diffuse field response of the ear. The notch in the frequency region above 6 kHz is a reflection of the standing wave null

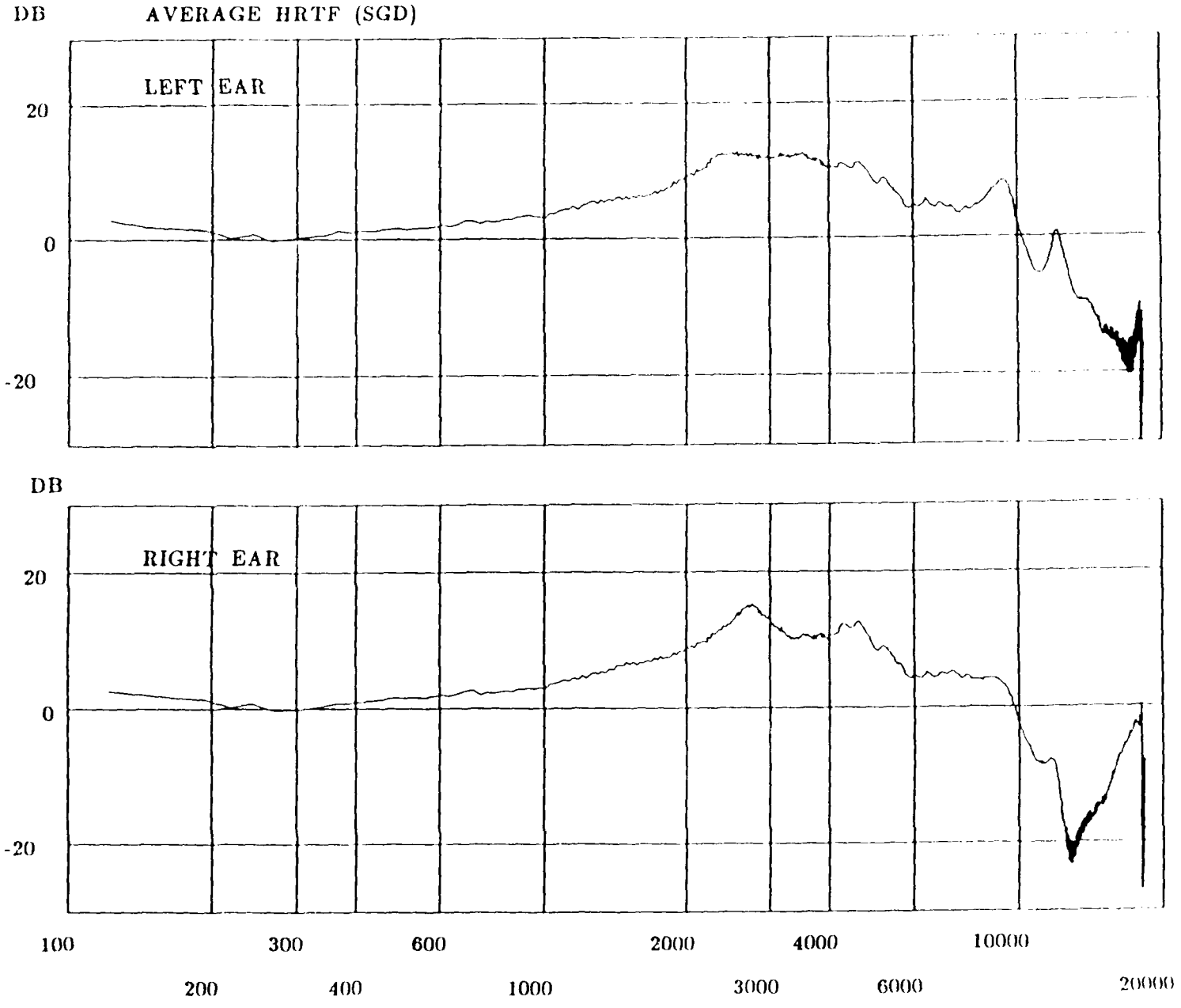


Figure 58. Same as Figure 57, except for Subject SGD

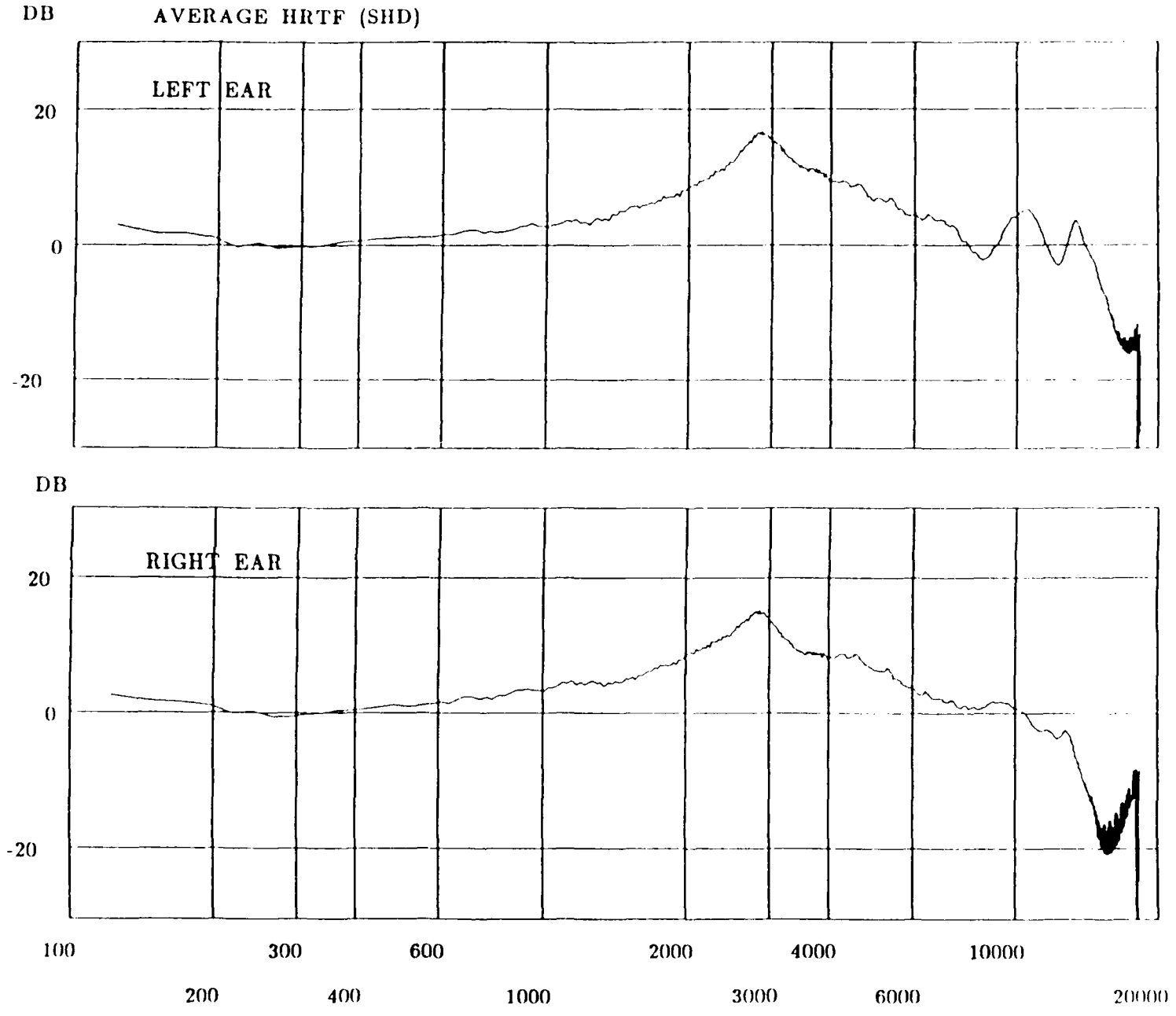


Figure 59. Same as Figure 57, except for Subject SHD

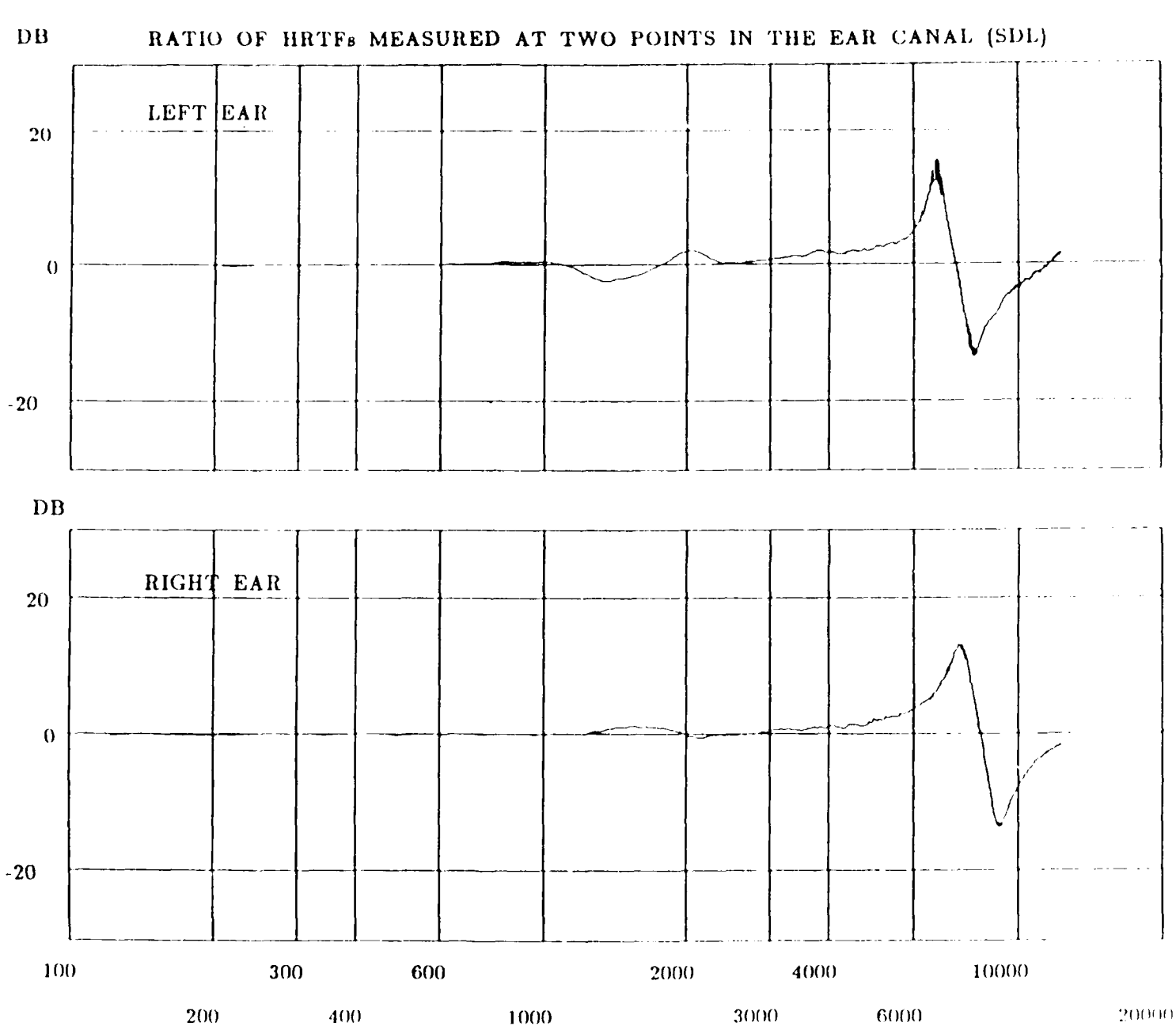


Figure 60. Ratio of HRTFs measured at two points (normal and extracted) in the ear canal for Subject SDL. The frequency at the bottom of the notch is used to estimate the distance from the probe tip to the eardrum when the microphone is in its normal position for HRTF measurement.

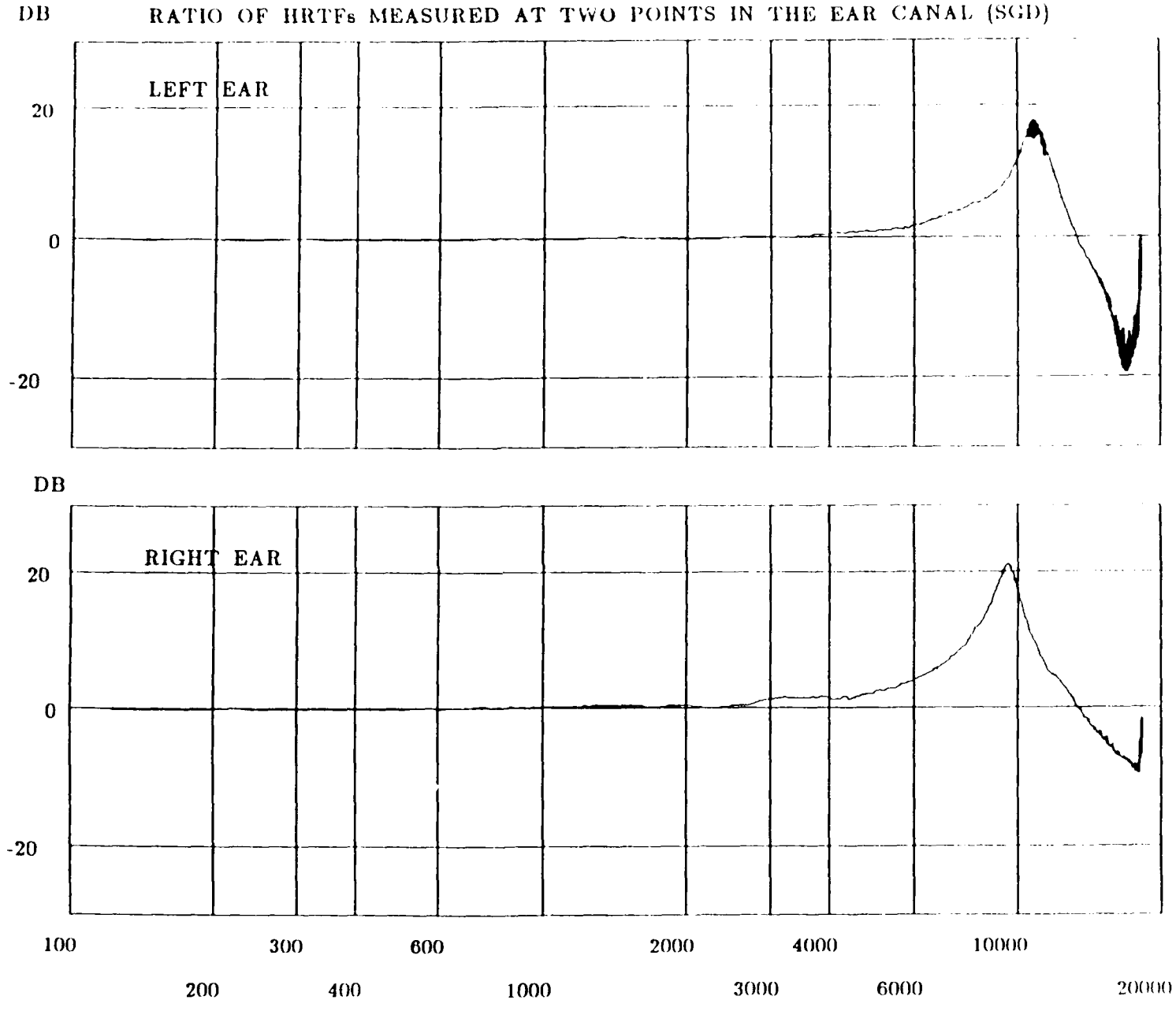


Figure 61. Same as Figure 60, except for subject SGD

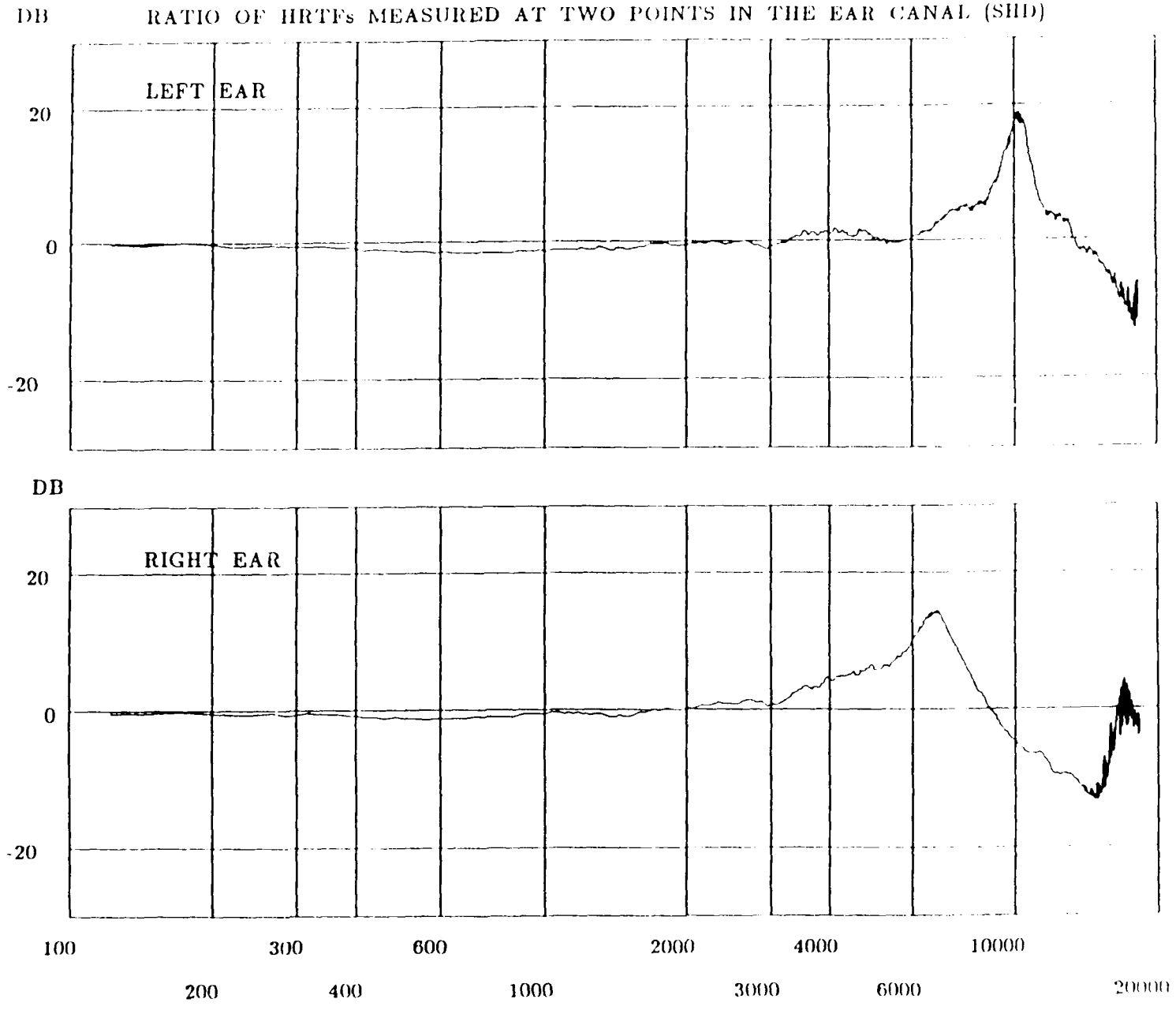


Figure 62. Same as Figure 60, except for subject SHD

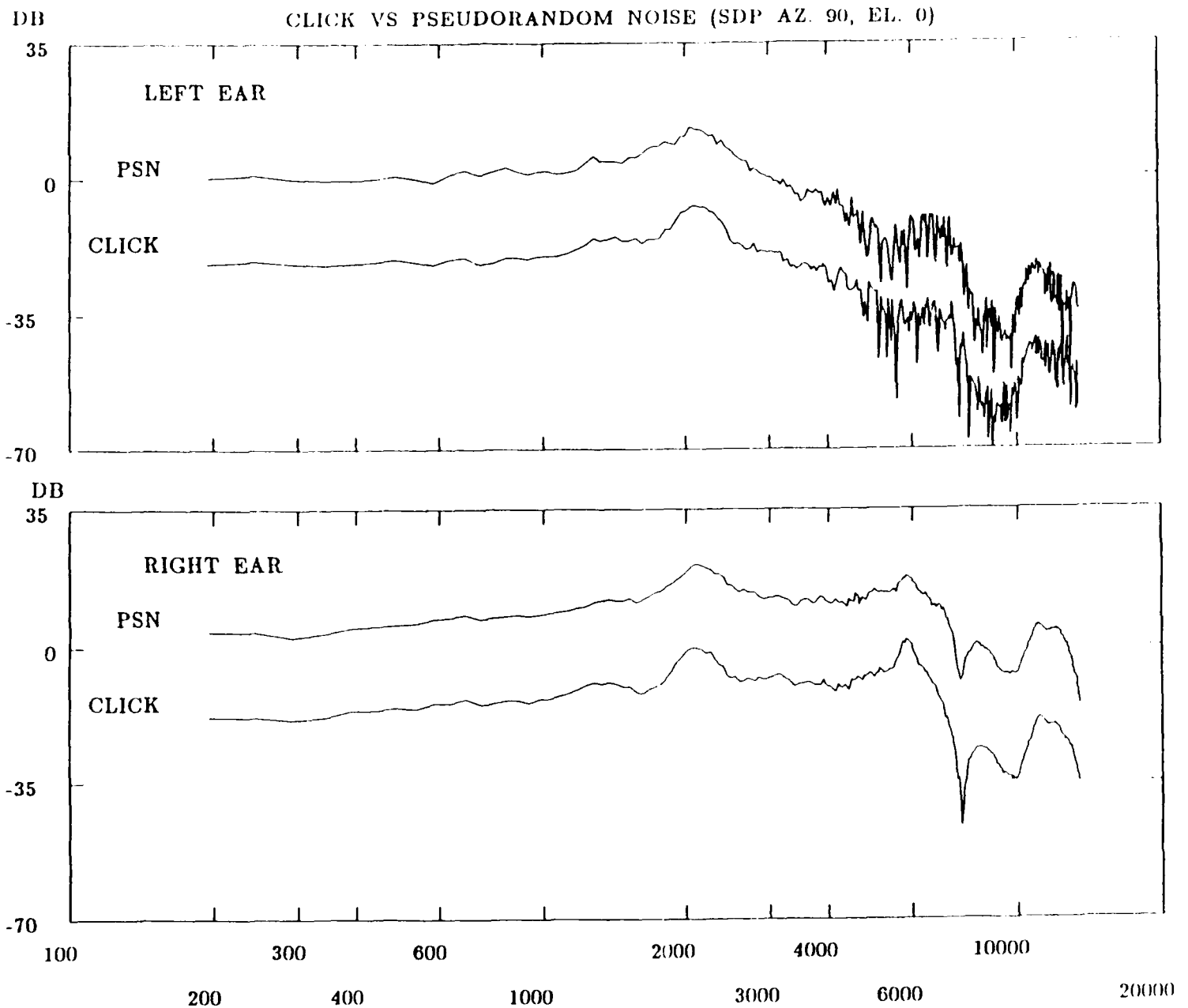


Figure 63. Magnitude of HRTFs measured at 90 degrees azimuth and 0 degrees elevation using a periodic pseudorandom noise and a click stimulus

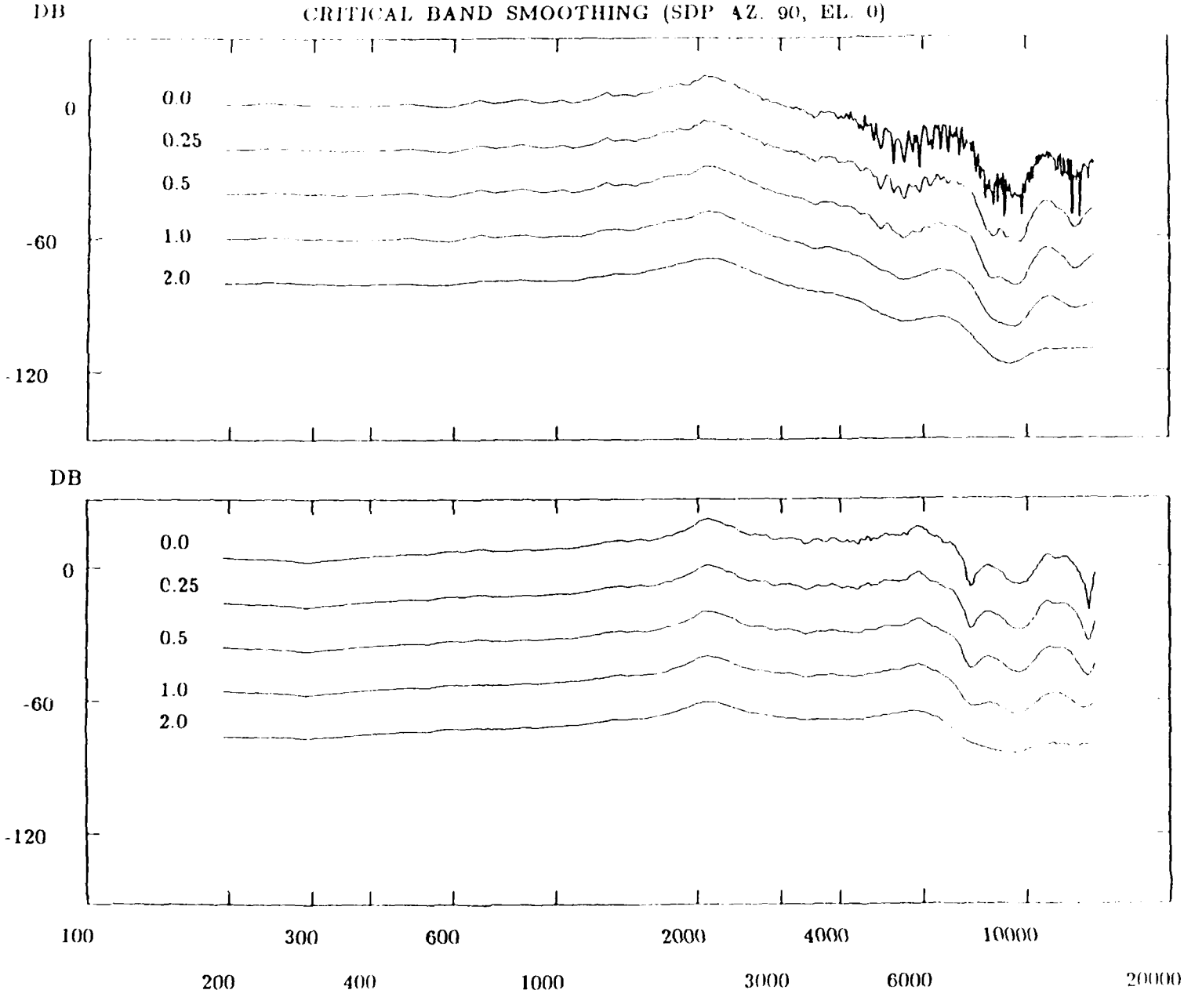


Figure 64. Unsmoothed magnitude of HRTF measured at 90 degrees azimuth and 0 degrees elevation and magnitudes that were smoothed using 0.25, 0.50, 1.0 and 2.0 critical bandwidths

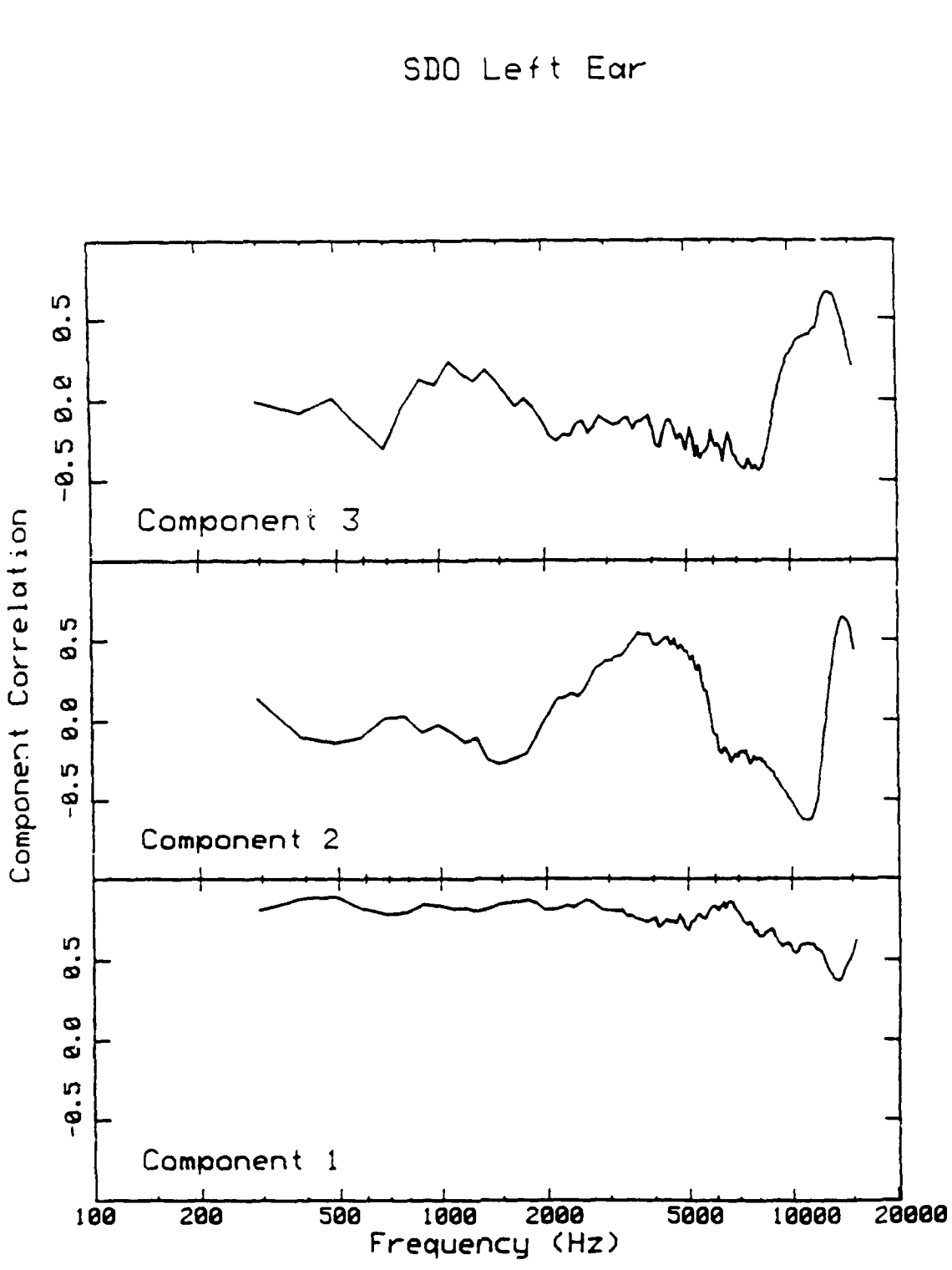


Figure 65. Correlations for the first three principal components fo · Subject SDO. The correlation indicates the contribution of a given frequency to the principal component

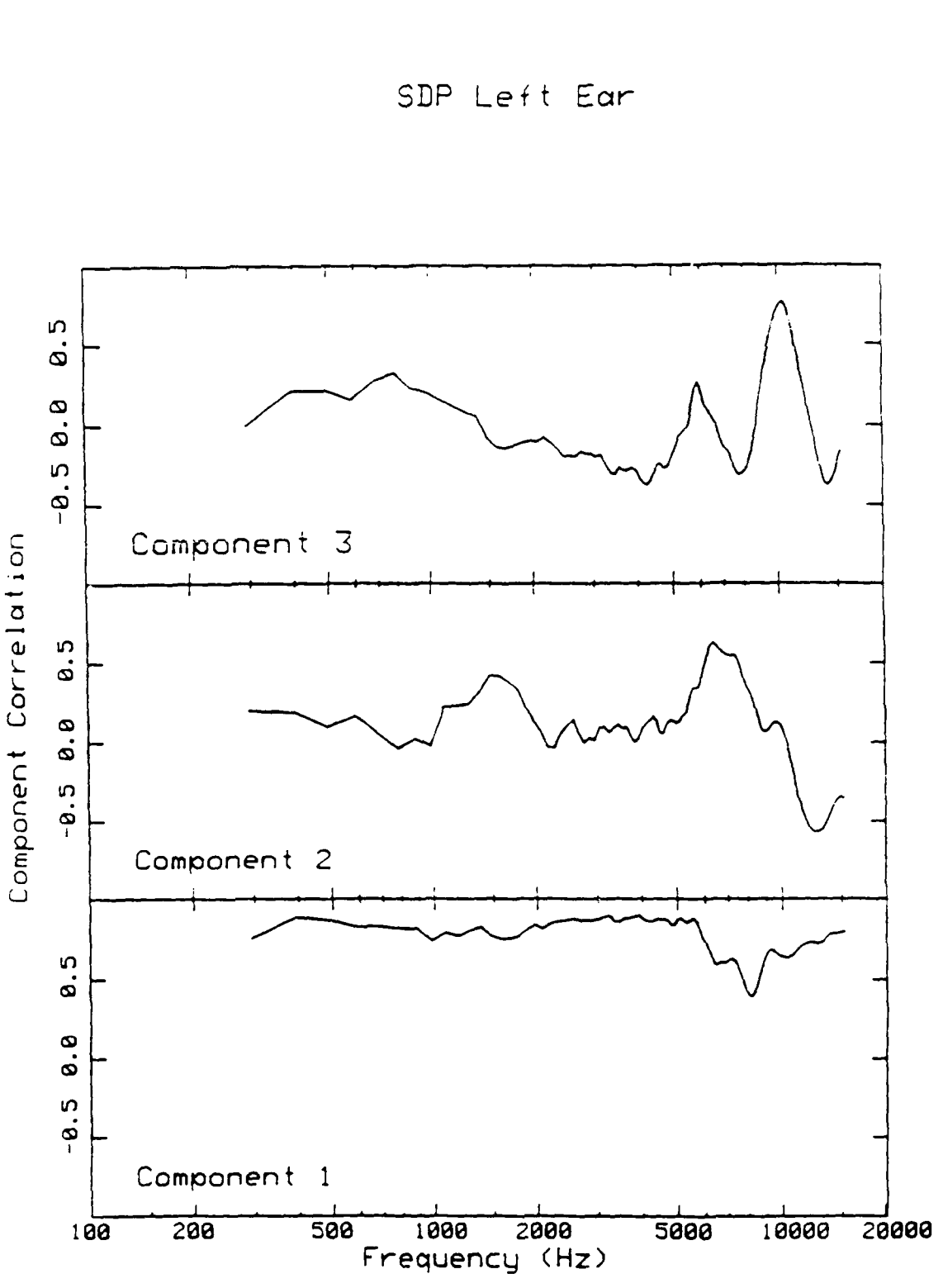


Figure 66. Same as Figure 65, except for Subject SDP

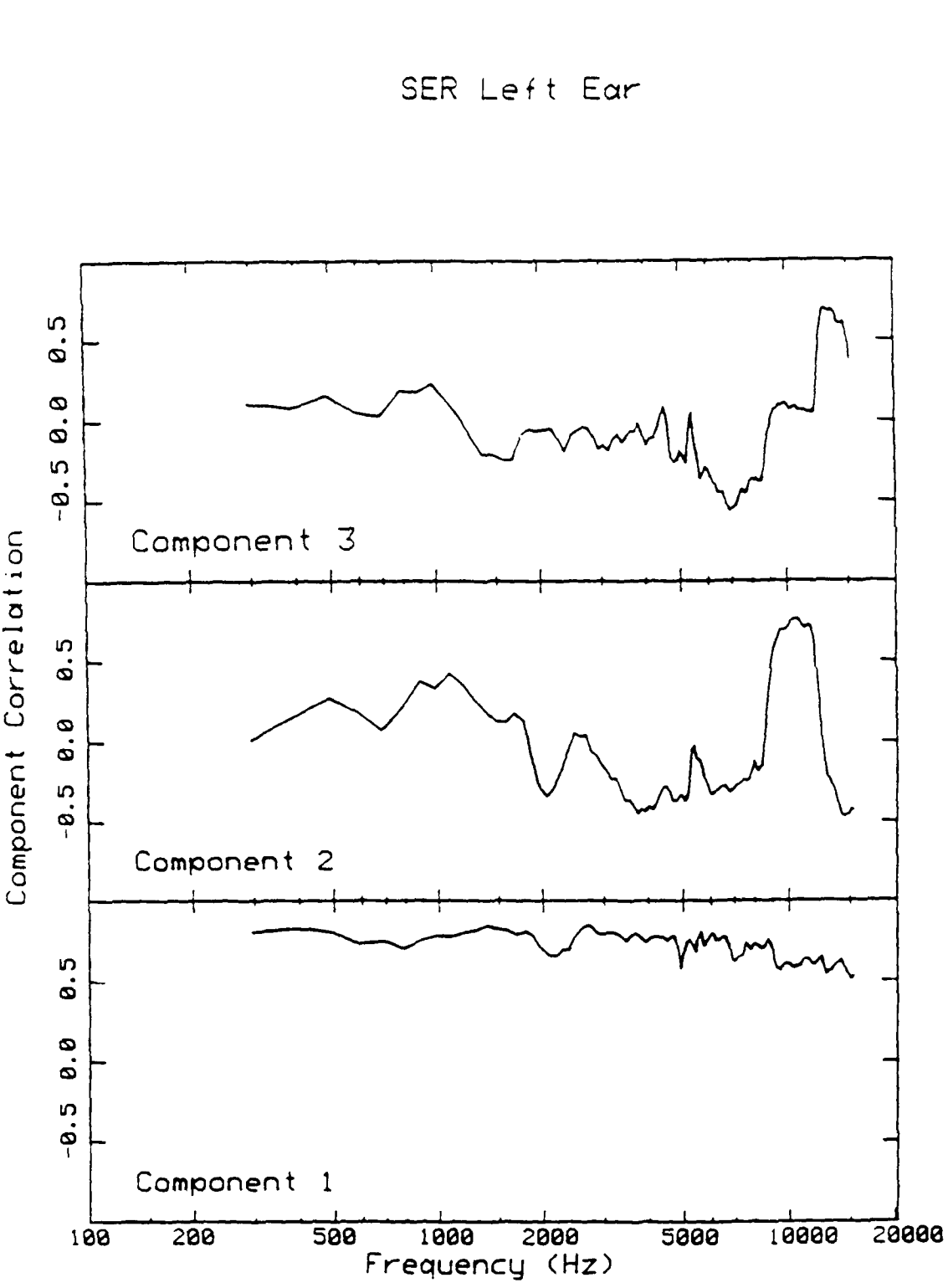


Figure 67. Same as Figure 65, except for Subject SER

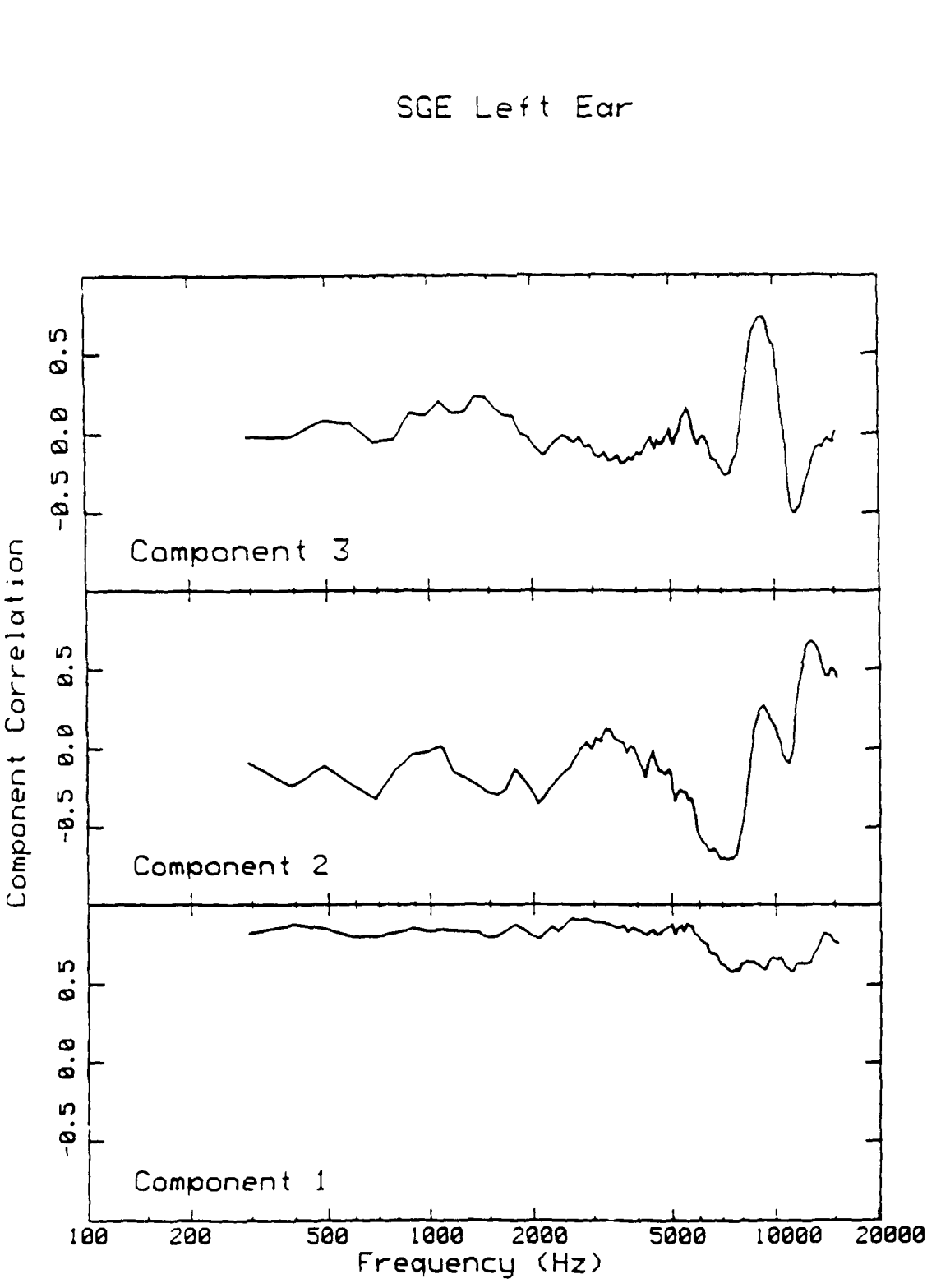


Figure 68. Same as Figure 65, except for Subject SGE

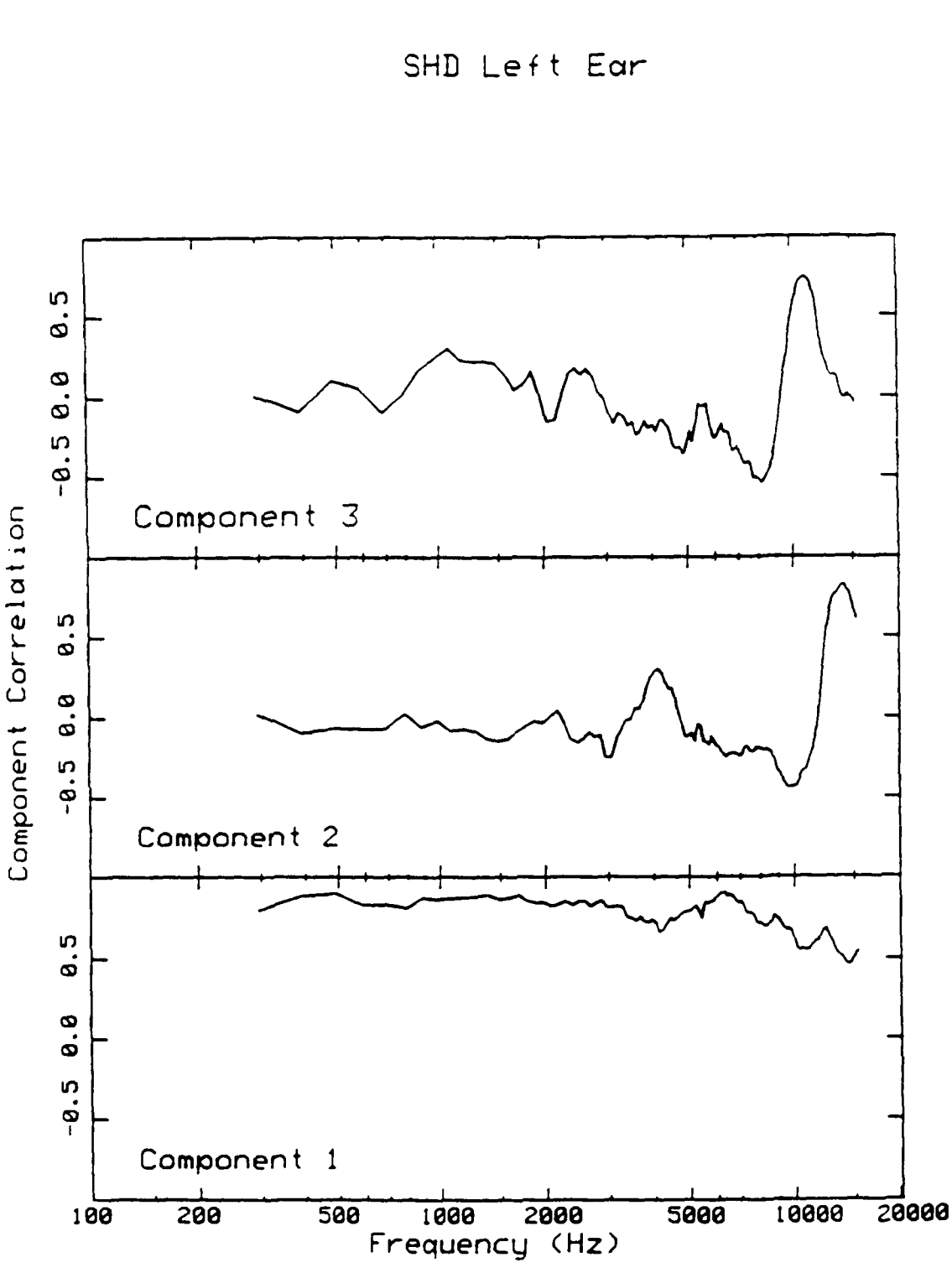


Figure 69. Same as Figure 65, except for Subject SHD

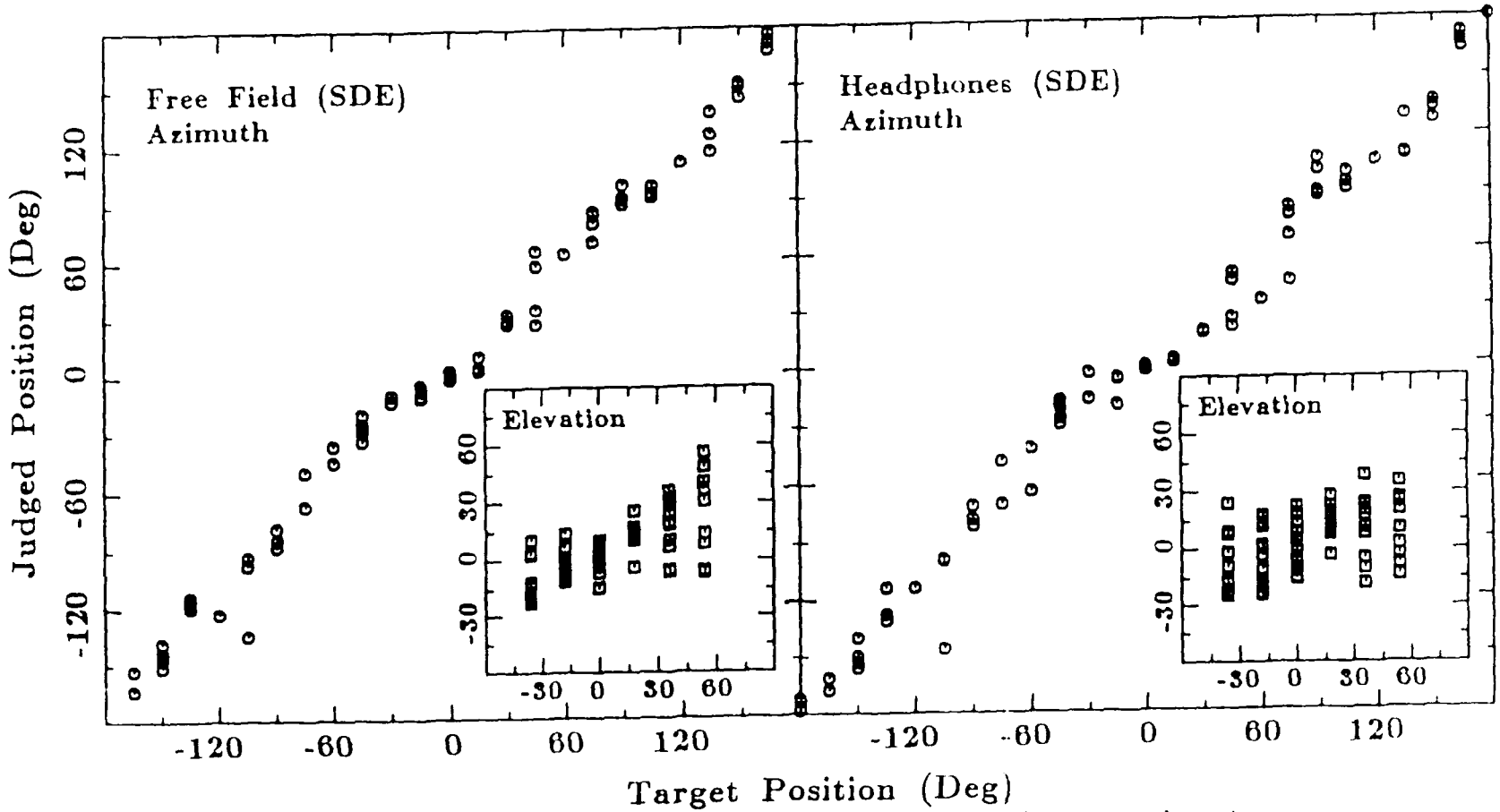


Figure 70. Scatterplots of actual source azimuth (and, in the insets,elevation) versus judged source azimuth for Subject SDE in both the free-field and headphone conditions. Each data point represents the centroid of at least 8 judgements. All 72 source positions are represented in each panel. Thus, data from 6 different source elevations are combined in the azimuth panels, and data from 24 different azimuths are combined in the elevation panels. Note that the scale is the same for azimuth and elevation plots.

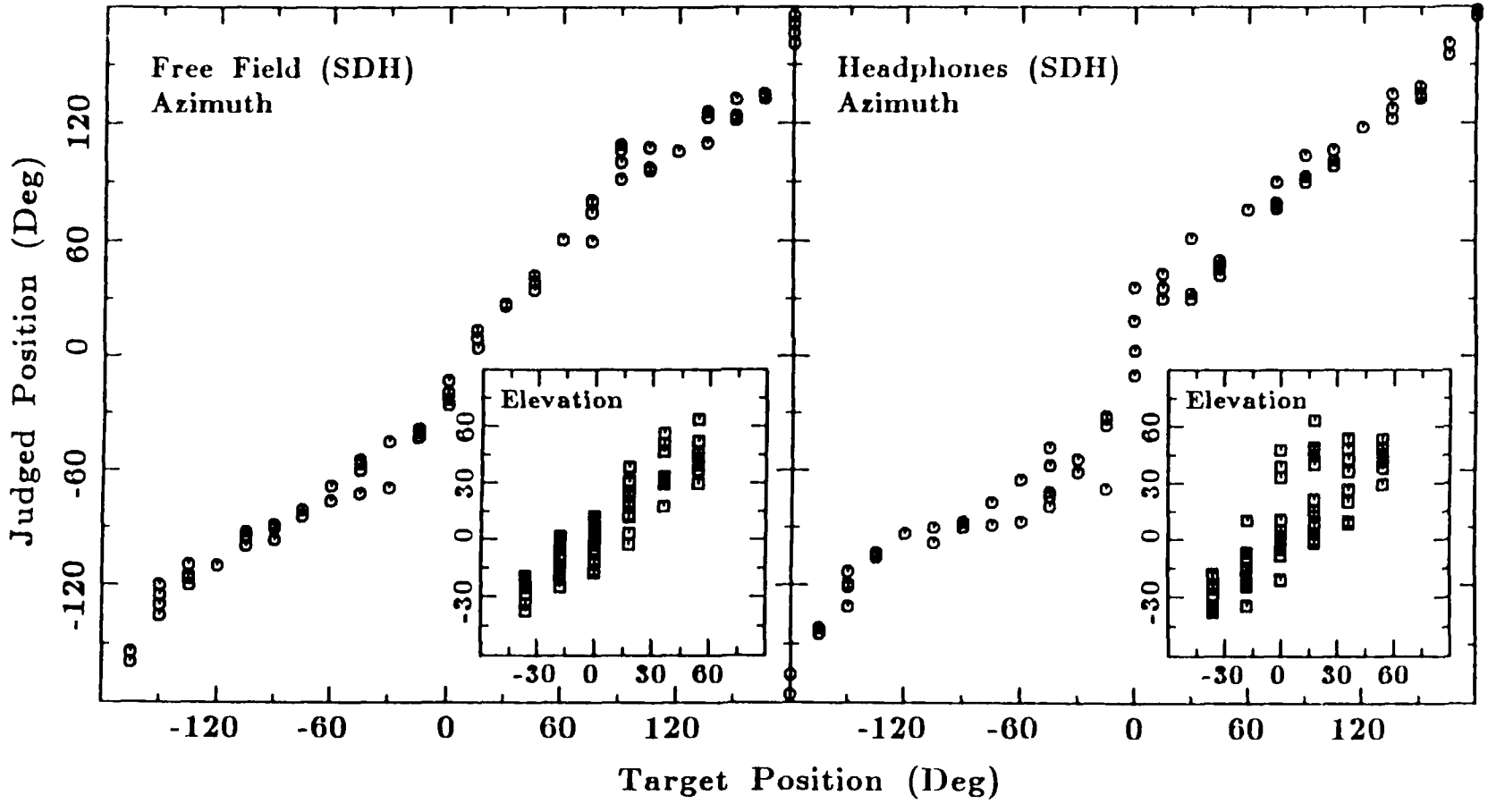


Figure 71. Same as Figure 70, except for Subject SDH

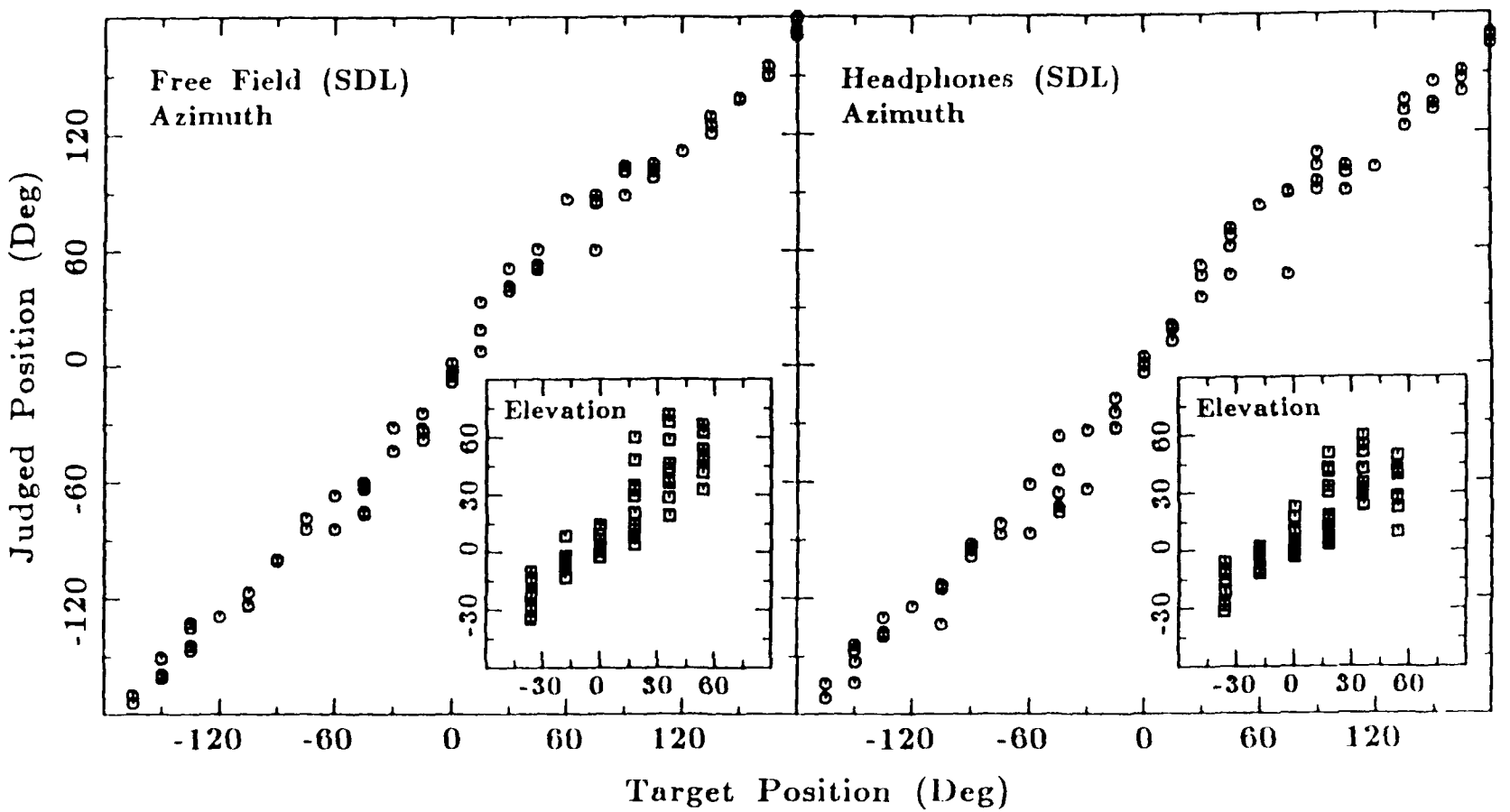


Figure 72. Same as Figure 70, except for Subject SDL

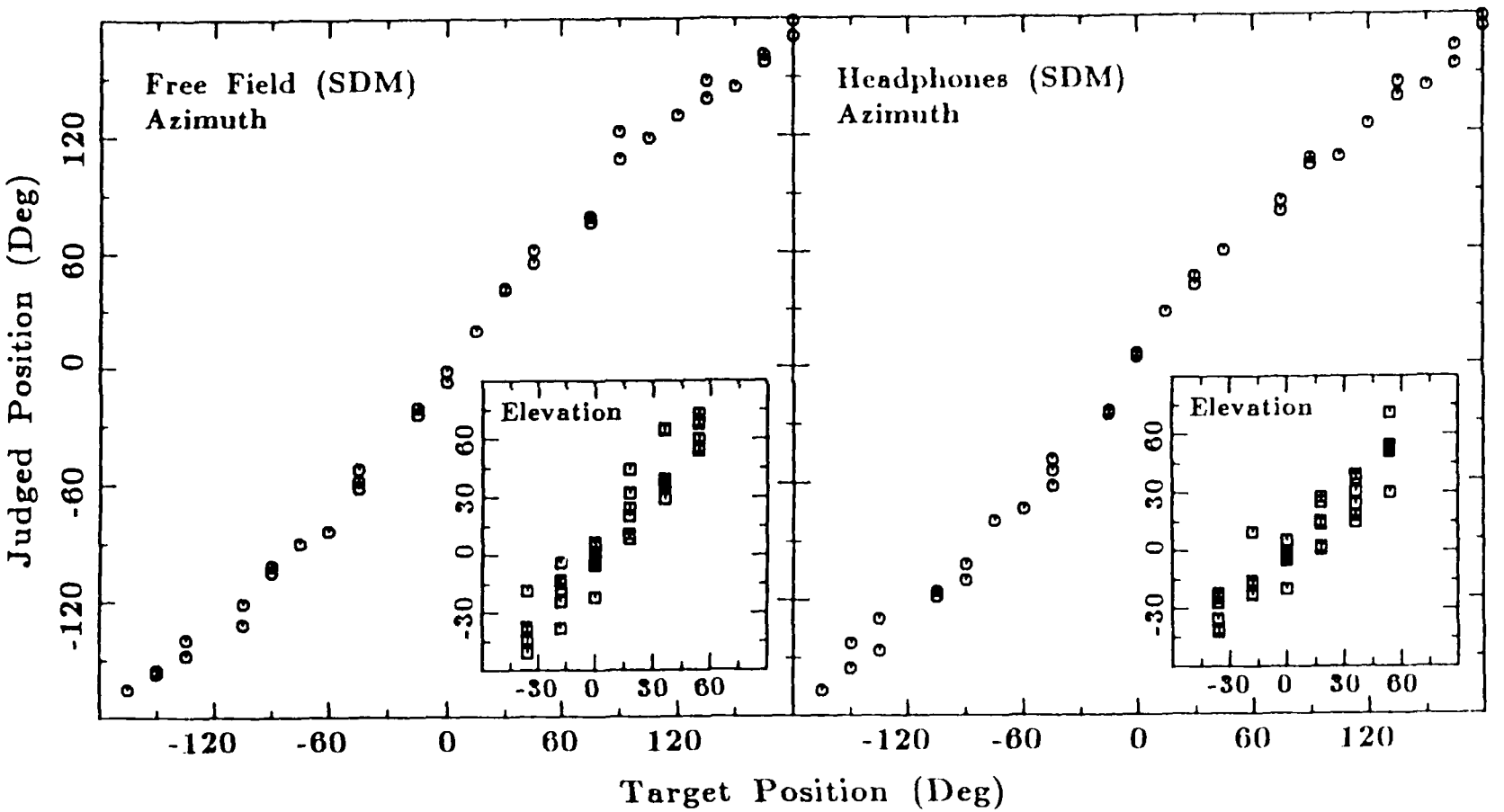


Figure 73. Same as Figure 70, except for Subject SDM

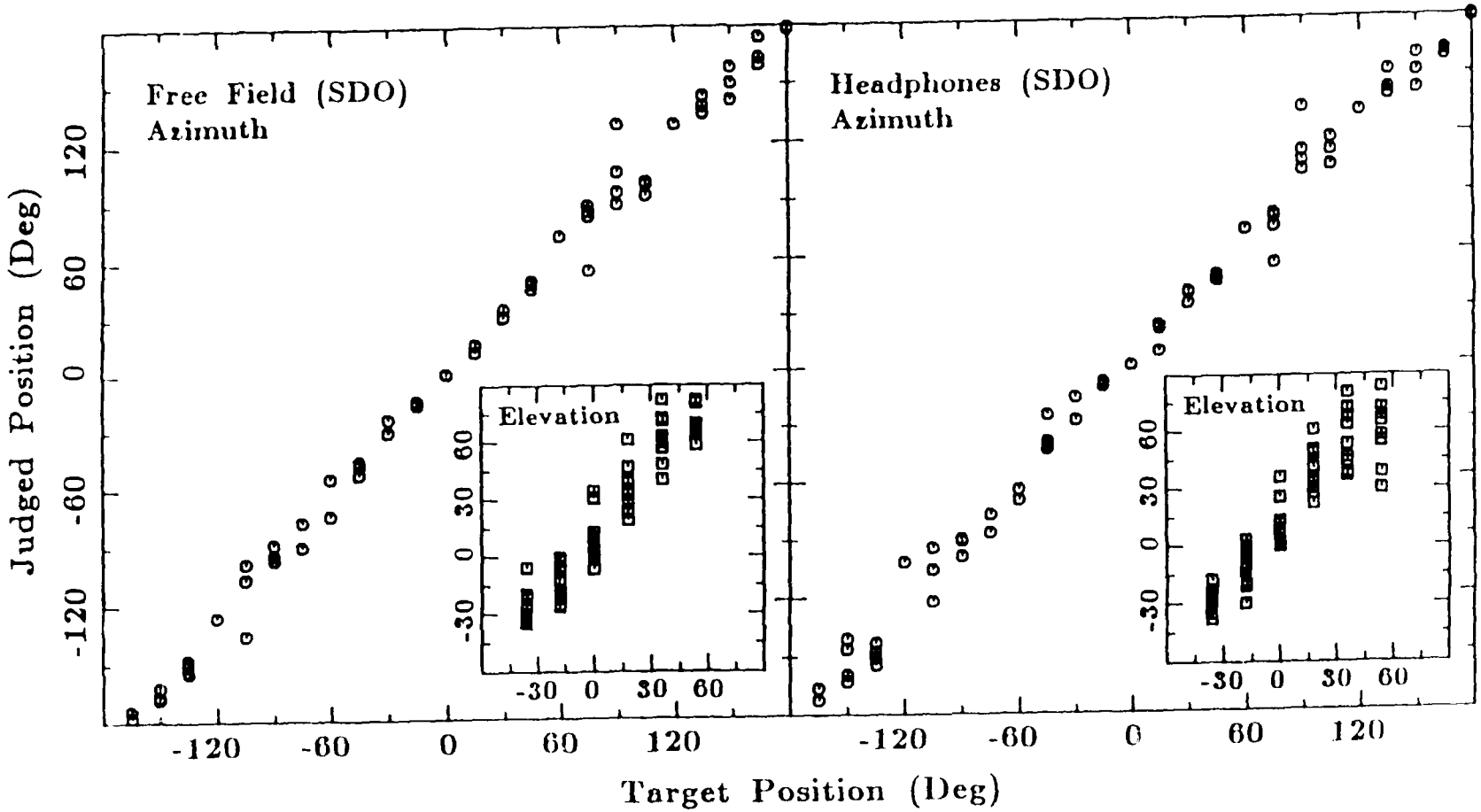


Figure 74. Same as Figure 70, except for Subject SDO

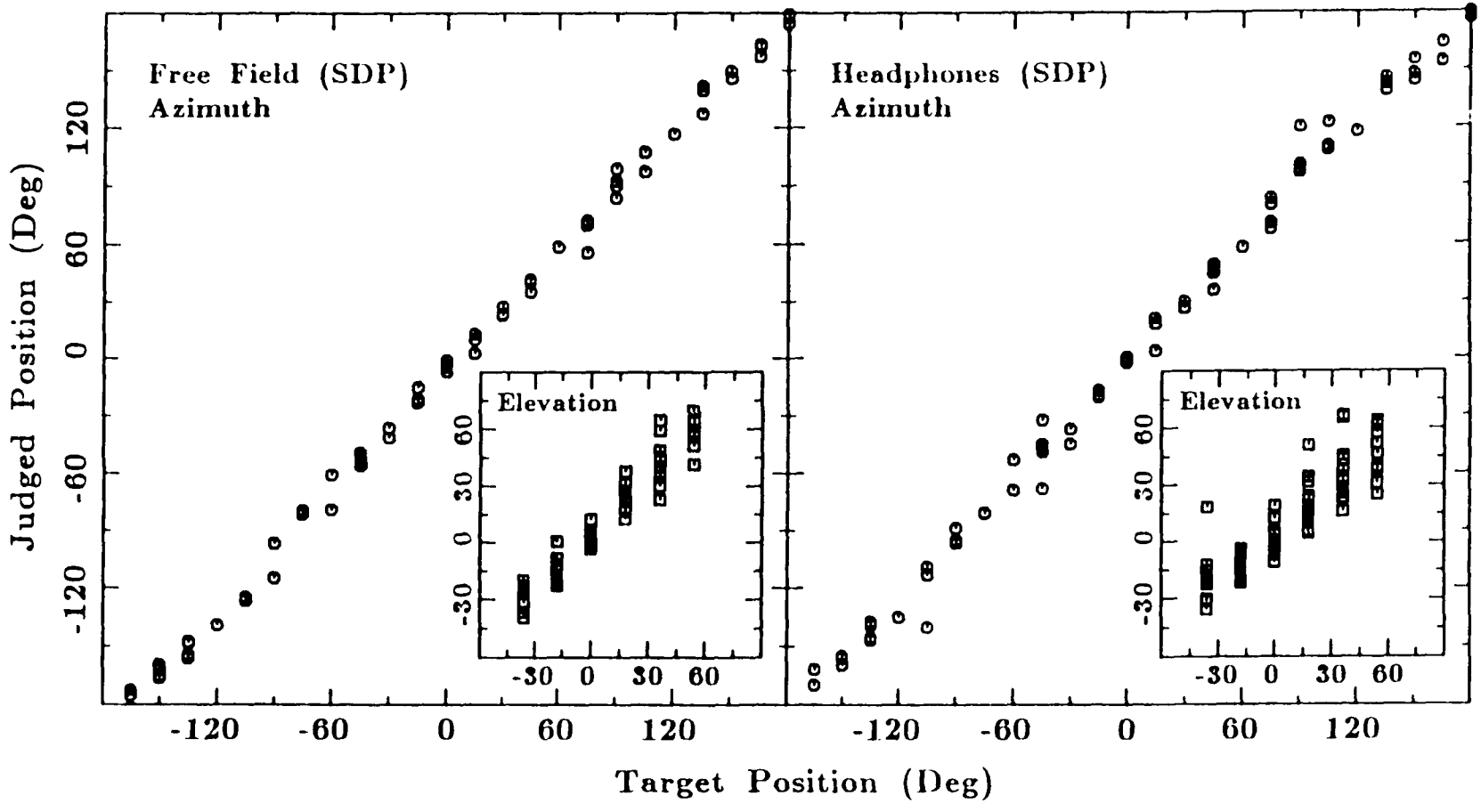


Figure 75. Same as Figure 70, except for Subject SDP

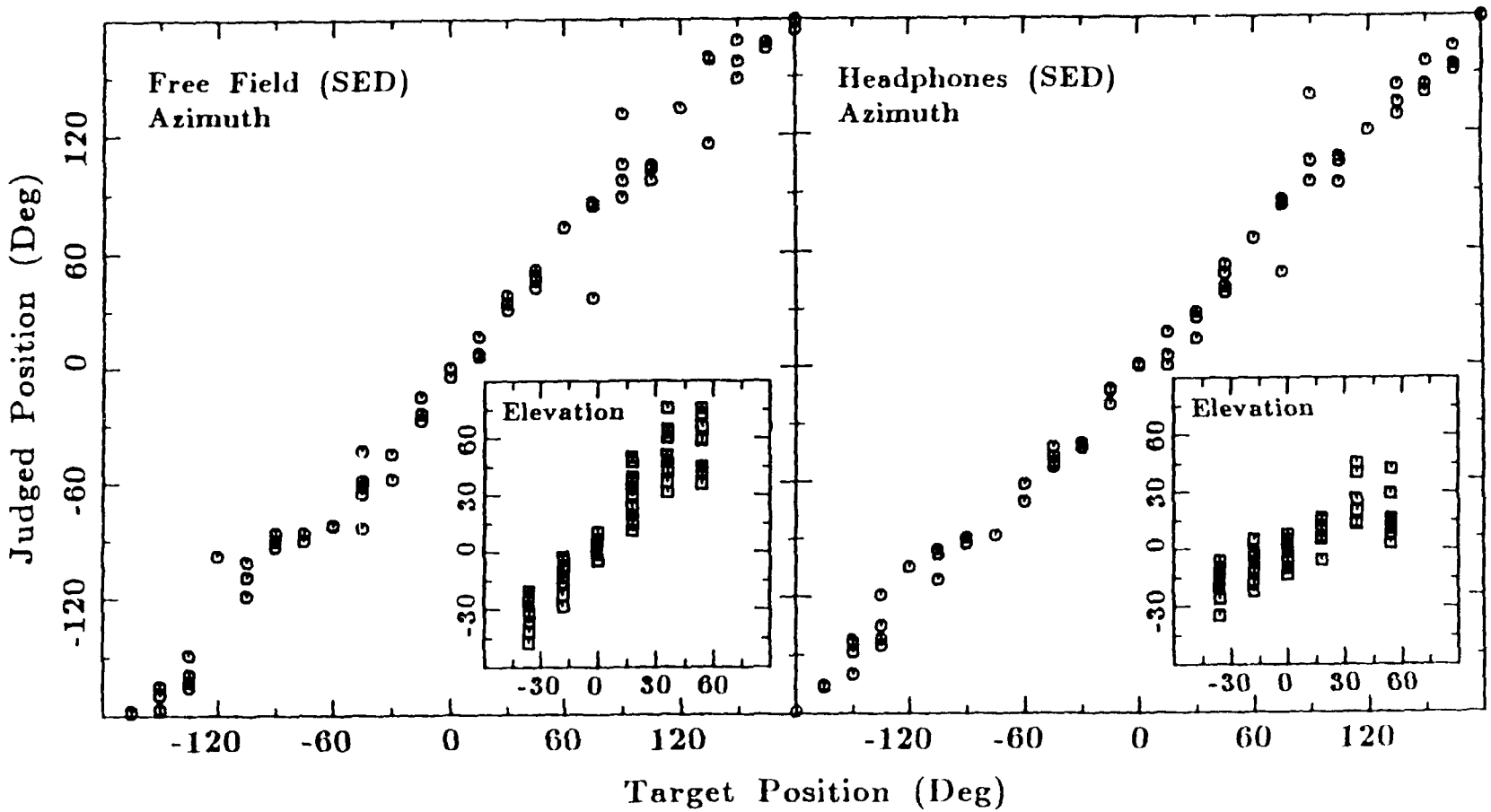


Figure 76. Same as Figure 70, except for Subject SED

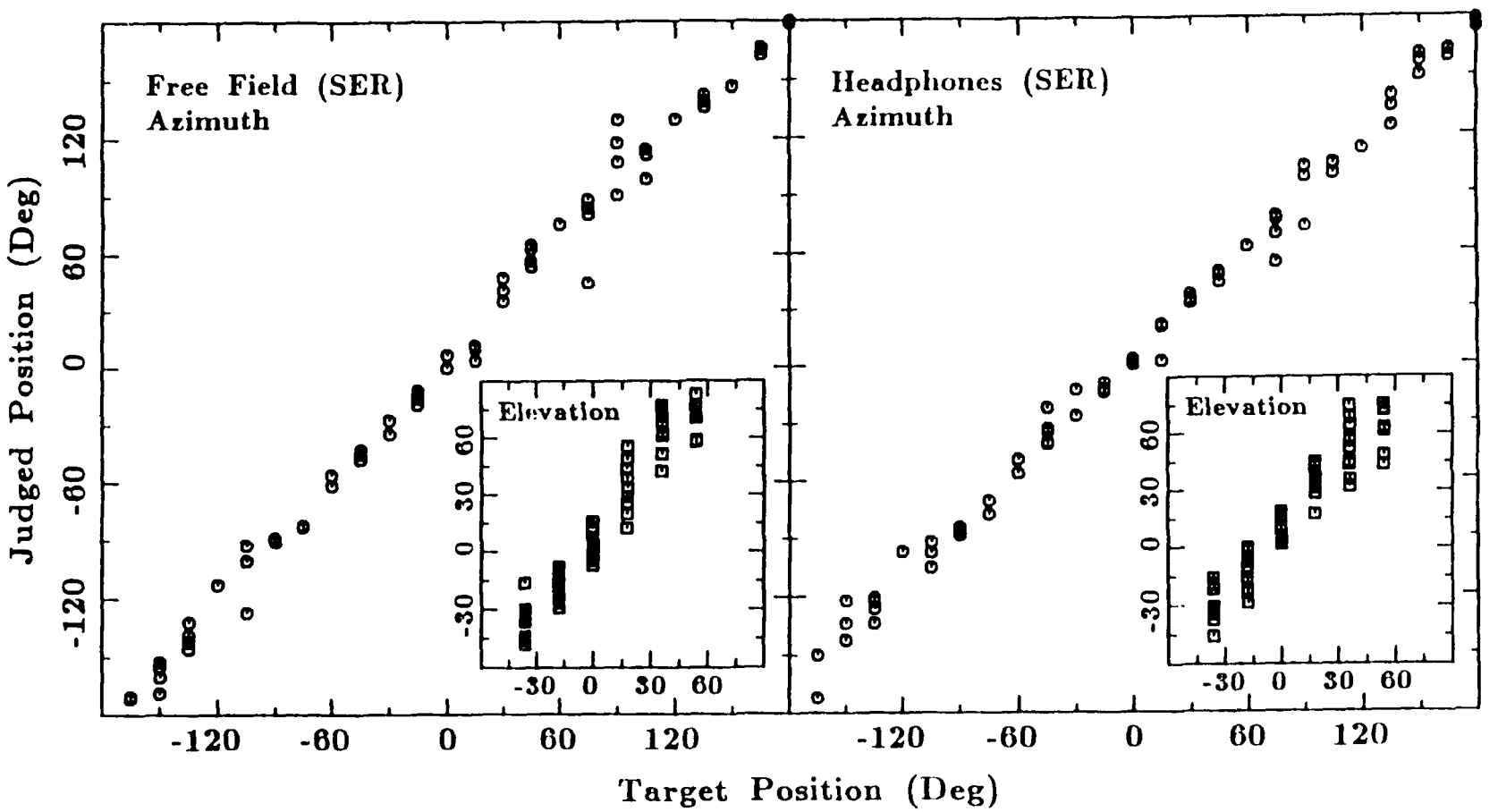


Figure 77. Same as Figure 70, except for Subject SER

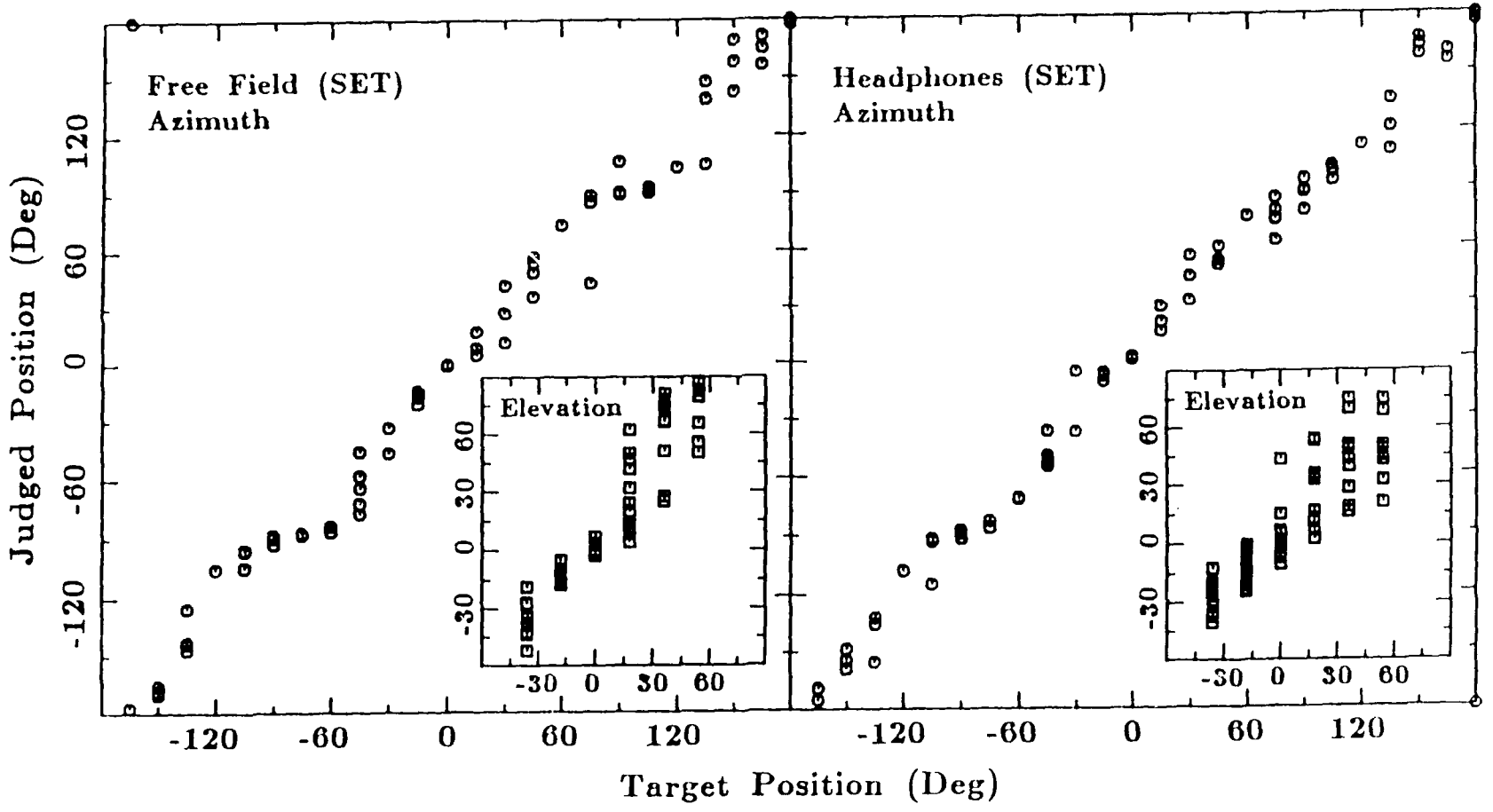


Figure 78. Same as Figure 70, except for Subject SET

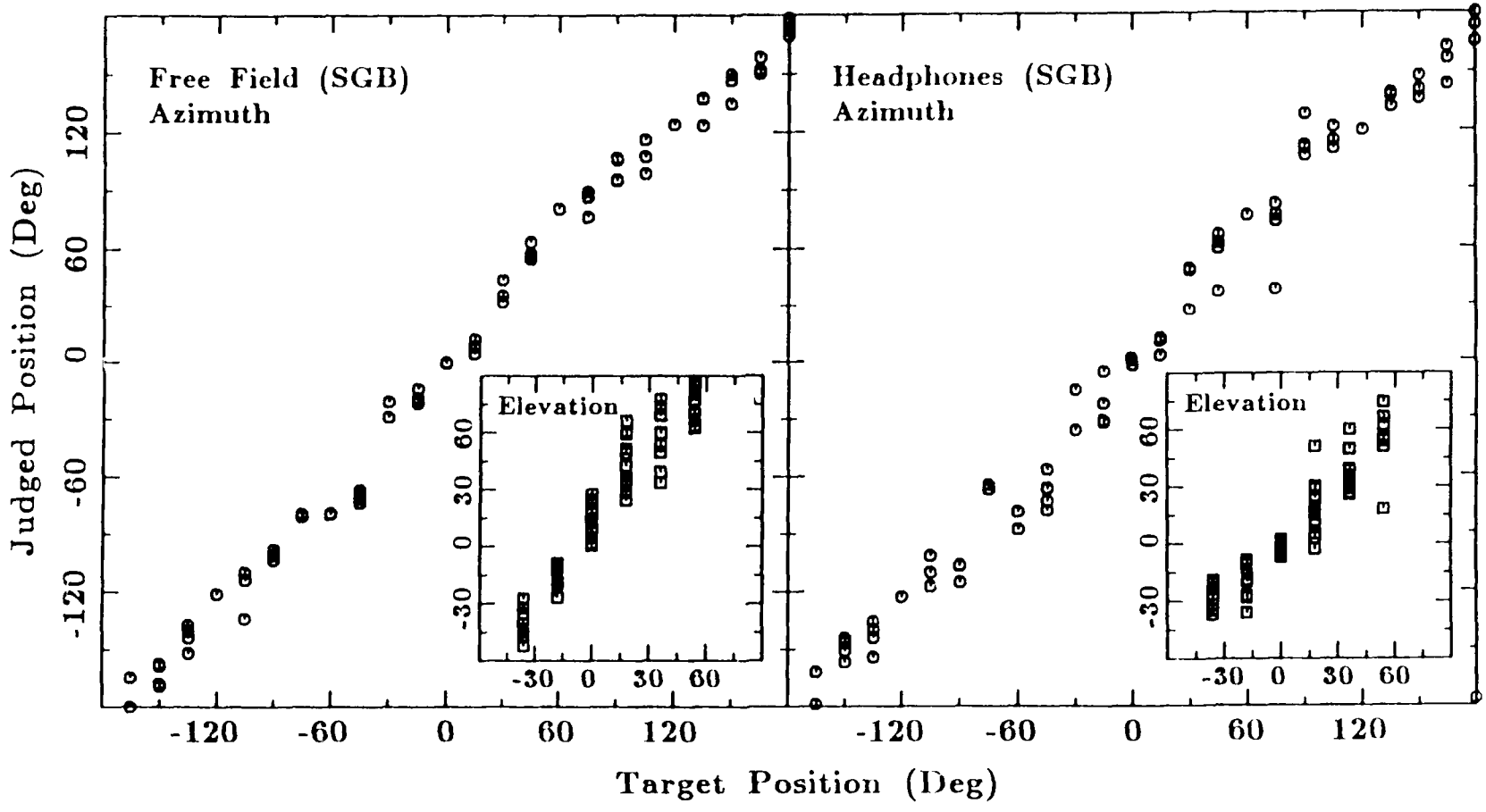


Figure 79. Same as Figure 70, except for Subject SGB

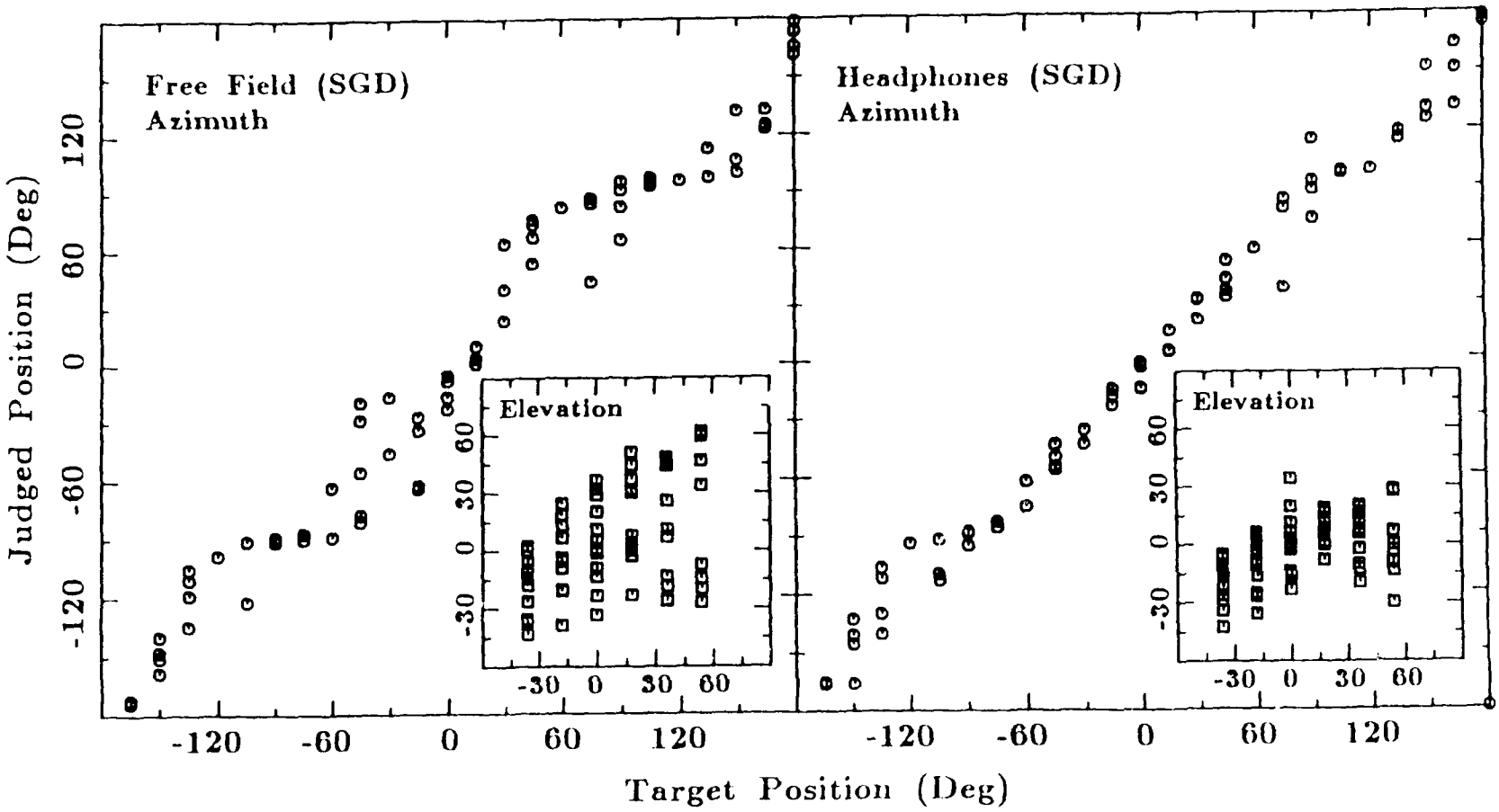


Figure 80. Same as Figure 70, except for Subject SGD

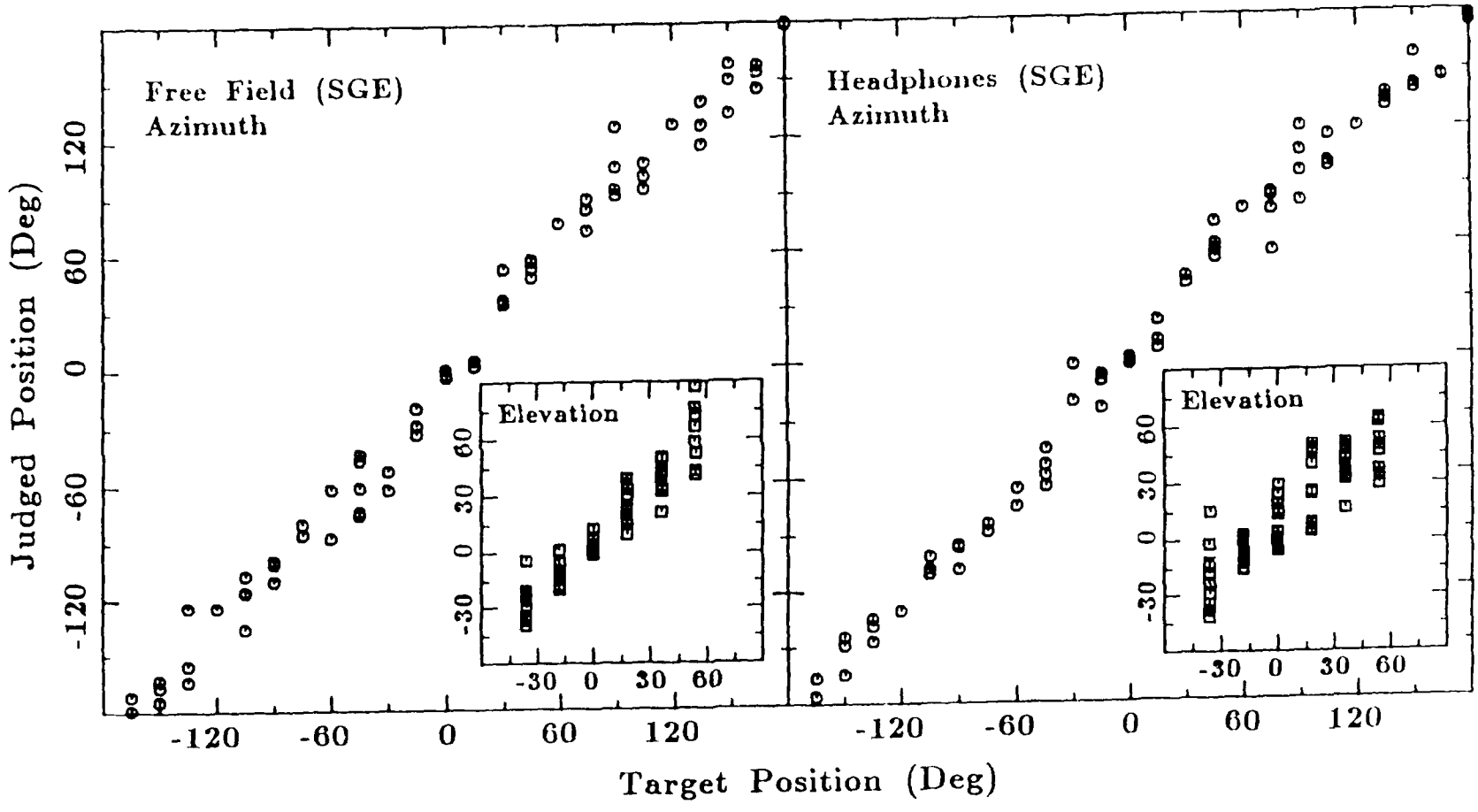


Figure 81. Same as Figure 70, except for Subject SGE

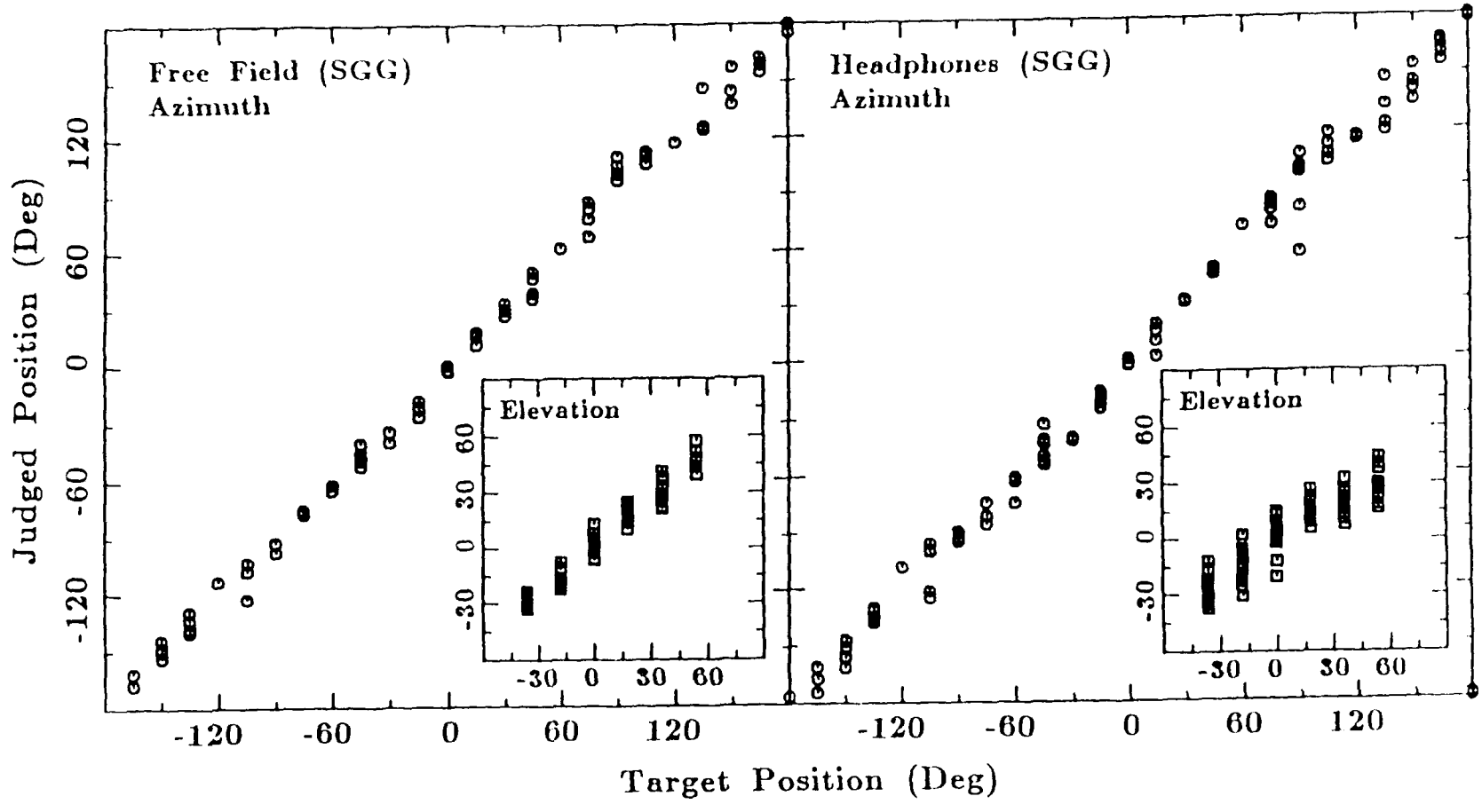


Figure 82. Same as Figure 70, except for Subject SGG

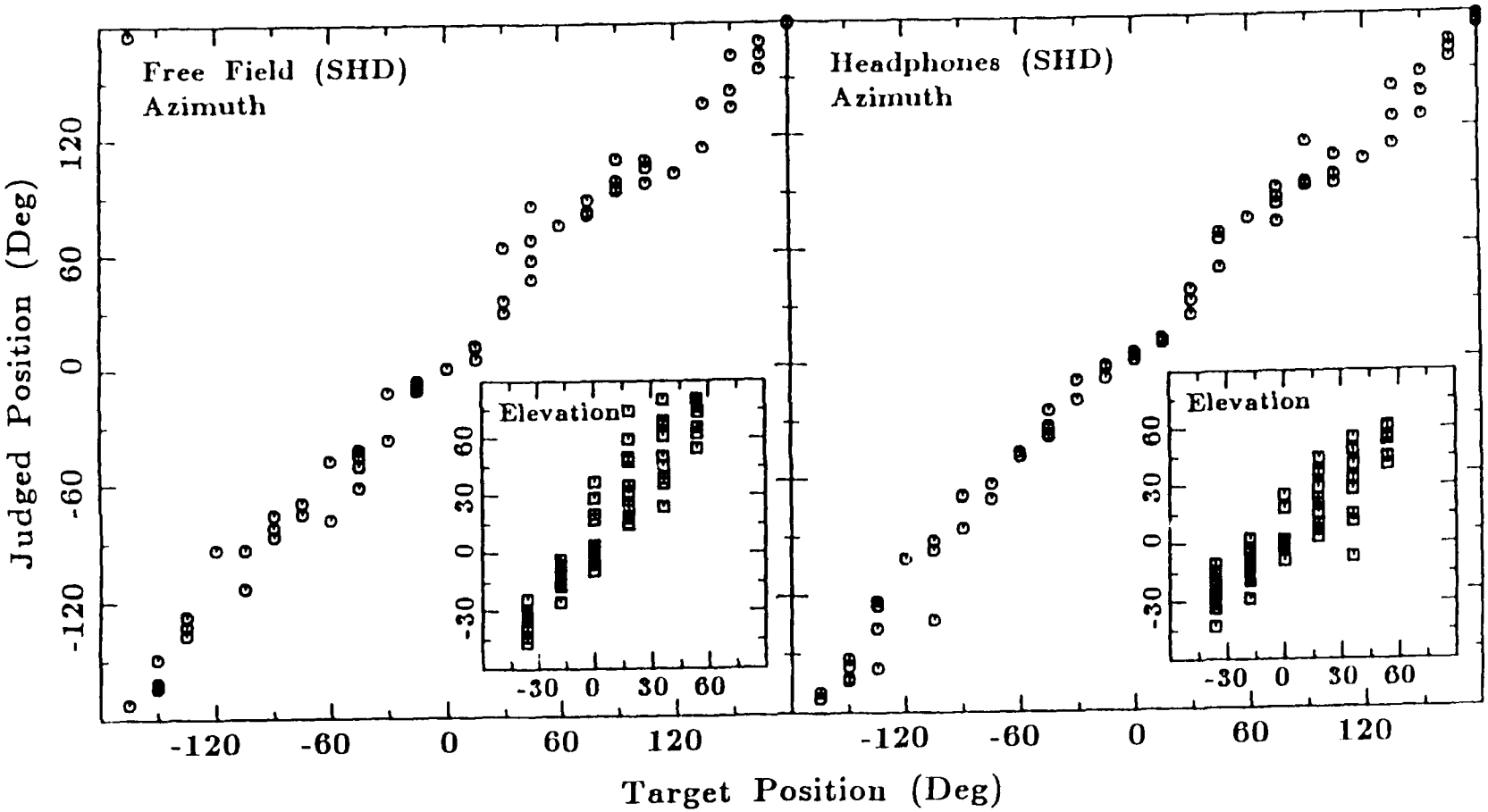


Figure 83. Same as Figure 70, except for Subject SHD

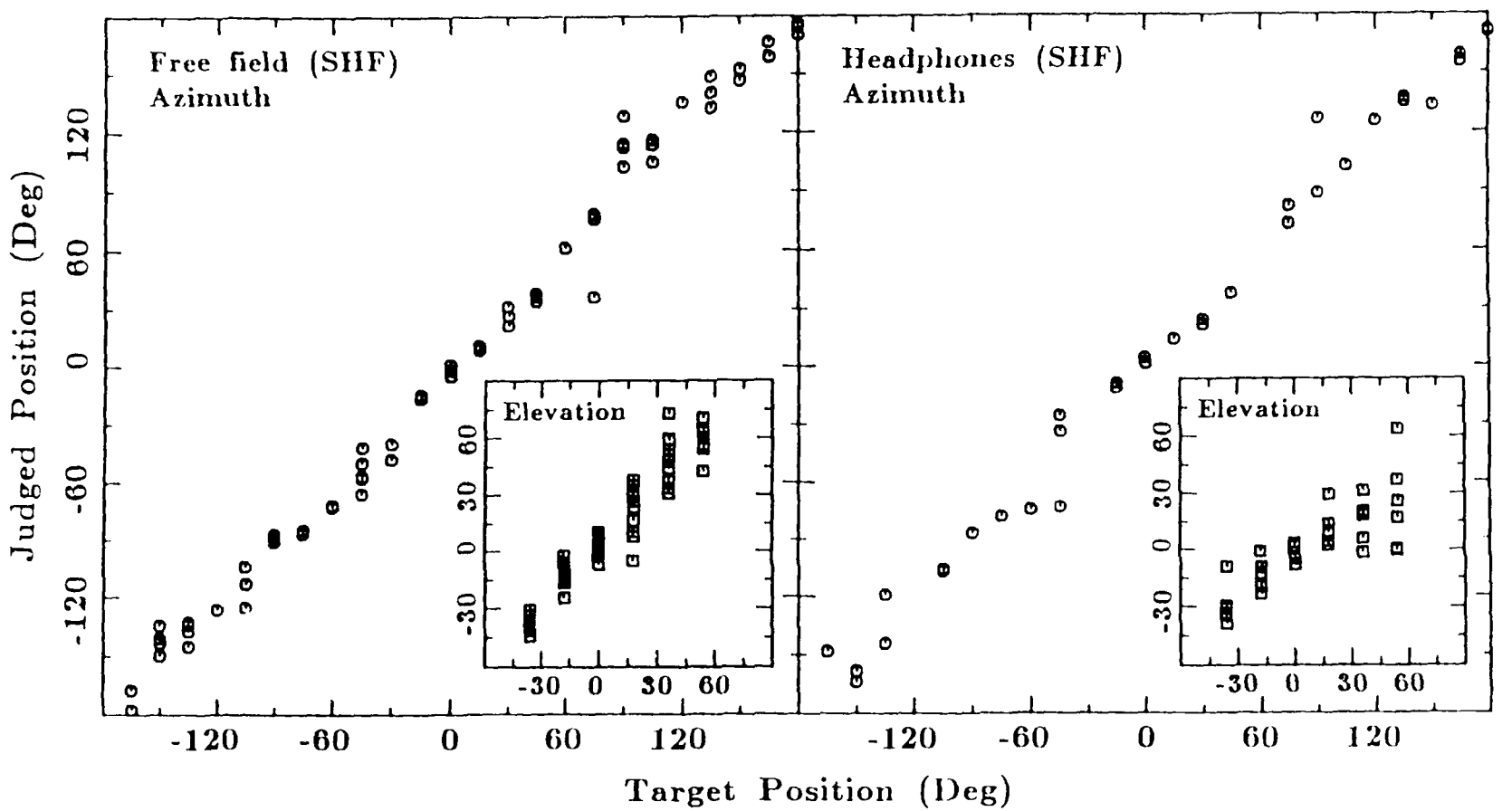


Figure 84. Same as Figure 70, except for Subject SHF

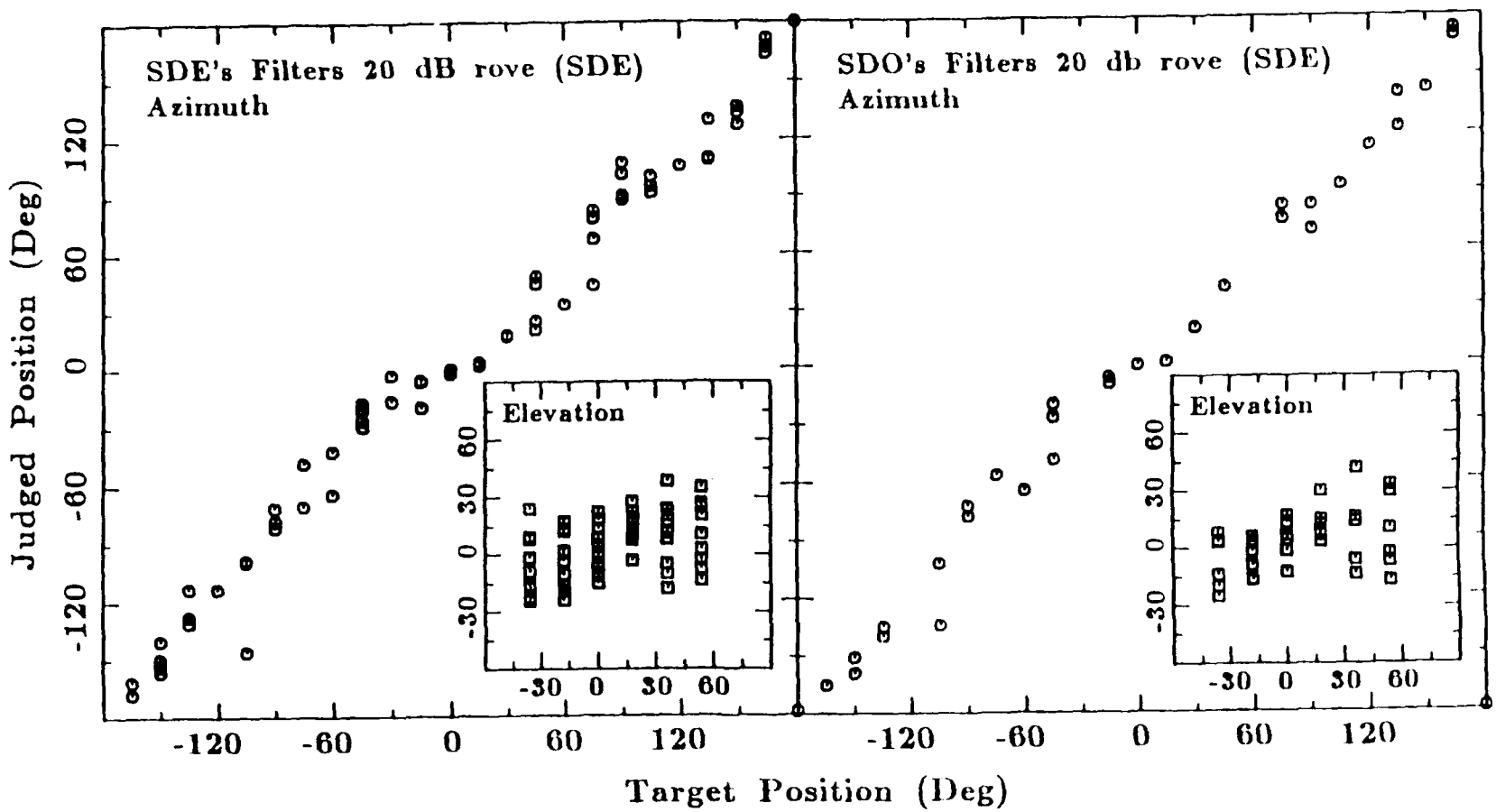


Figure 85. The right panel shows data from a condition in which Subject SDE localized 36 stimuli synthesized from HRTFs measured from Subject SDO. The format of the figure is the same as for Figure 70. The panel on the left is included for comparison and shows SDE's performance with stimuli based on SDE's own HRTFs

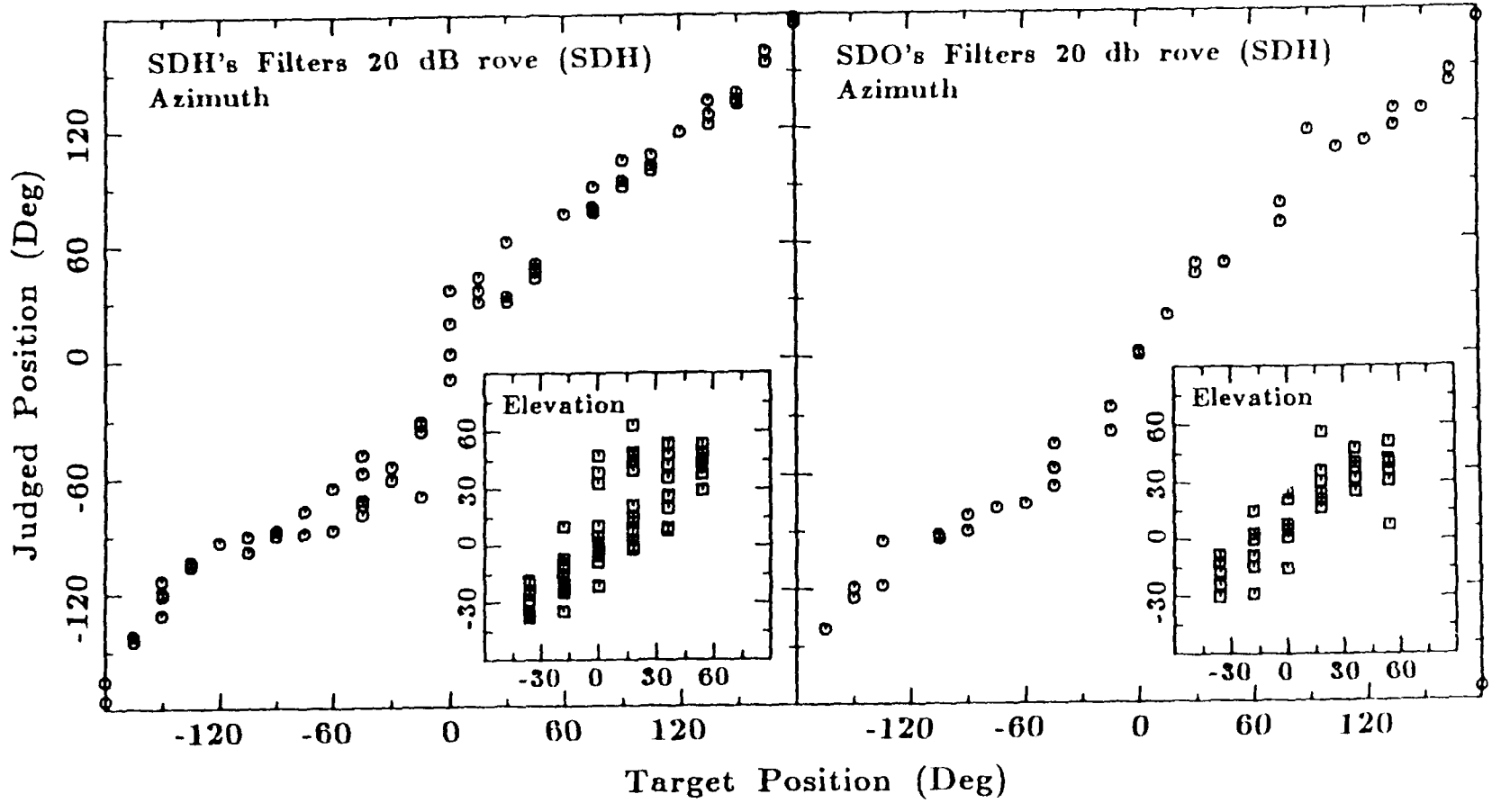


Figure 86. Same as Figure 85, except for Subject SDH

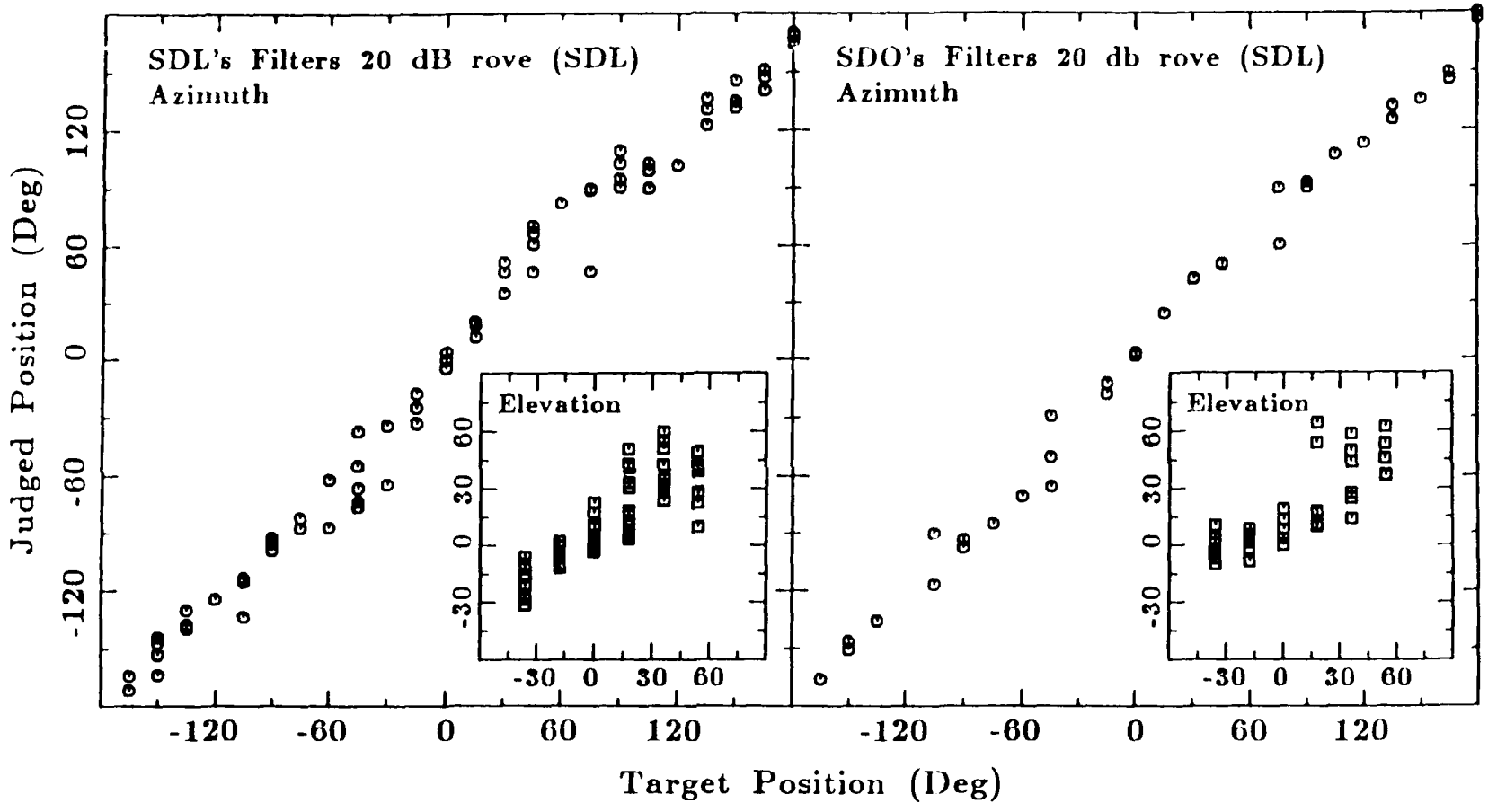


Figure 87. Same as Figure 85, except for Subject SDL

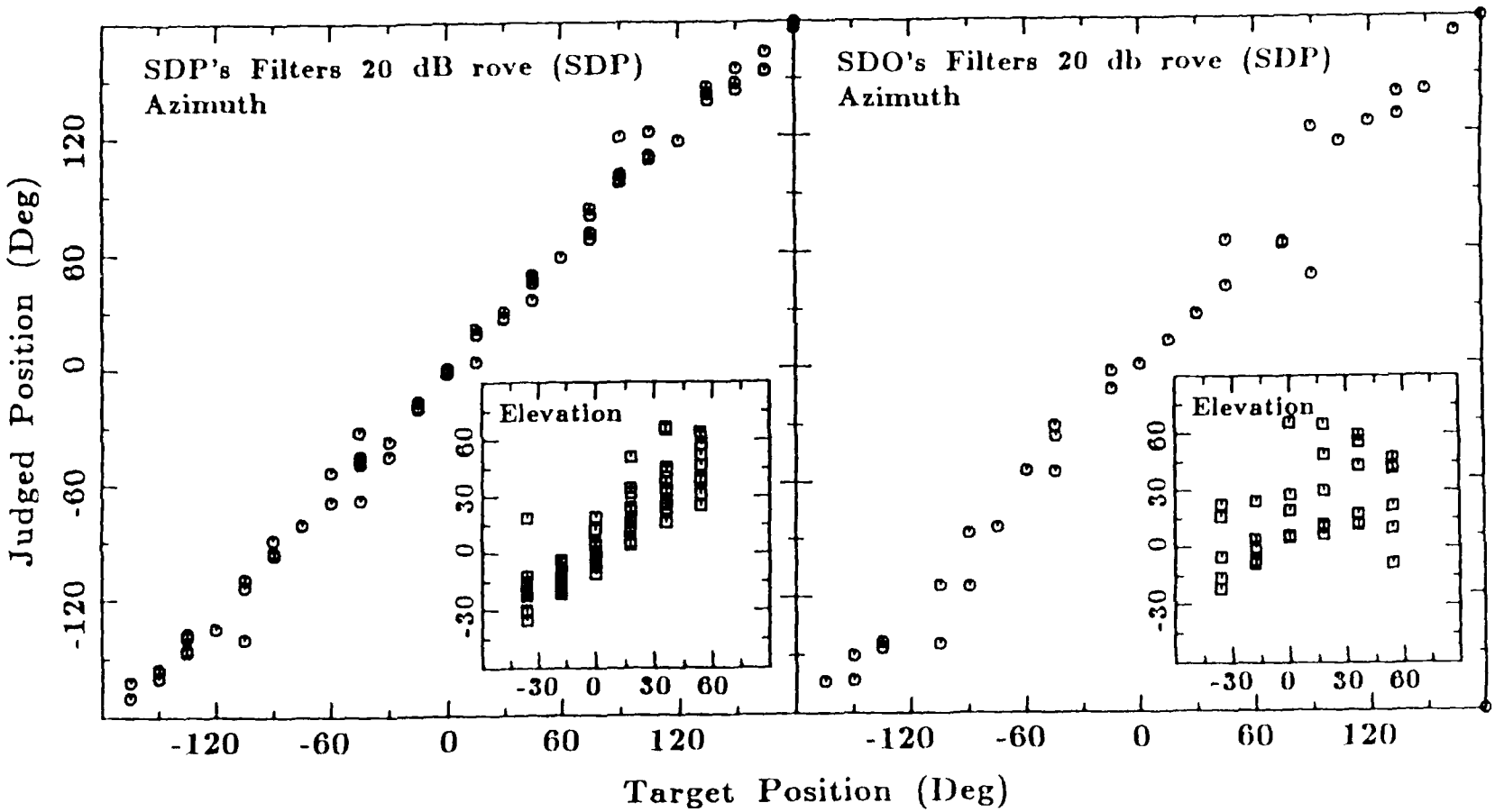


Figure 88. Same as Figure 85, except for Subject SDP

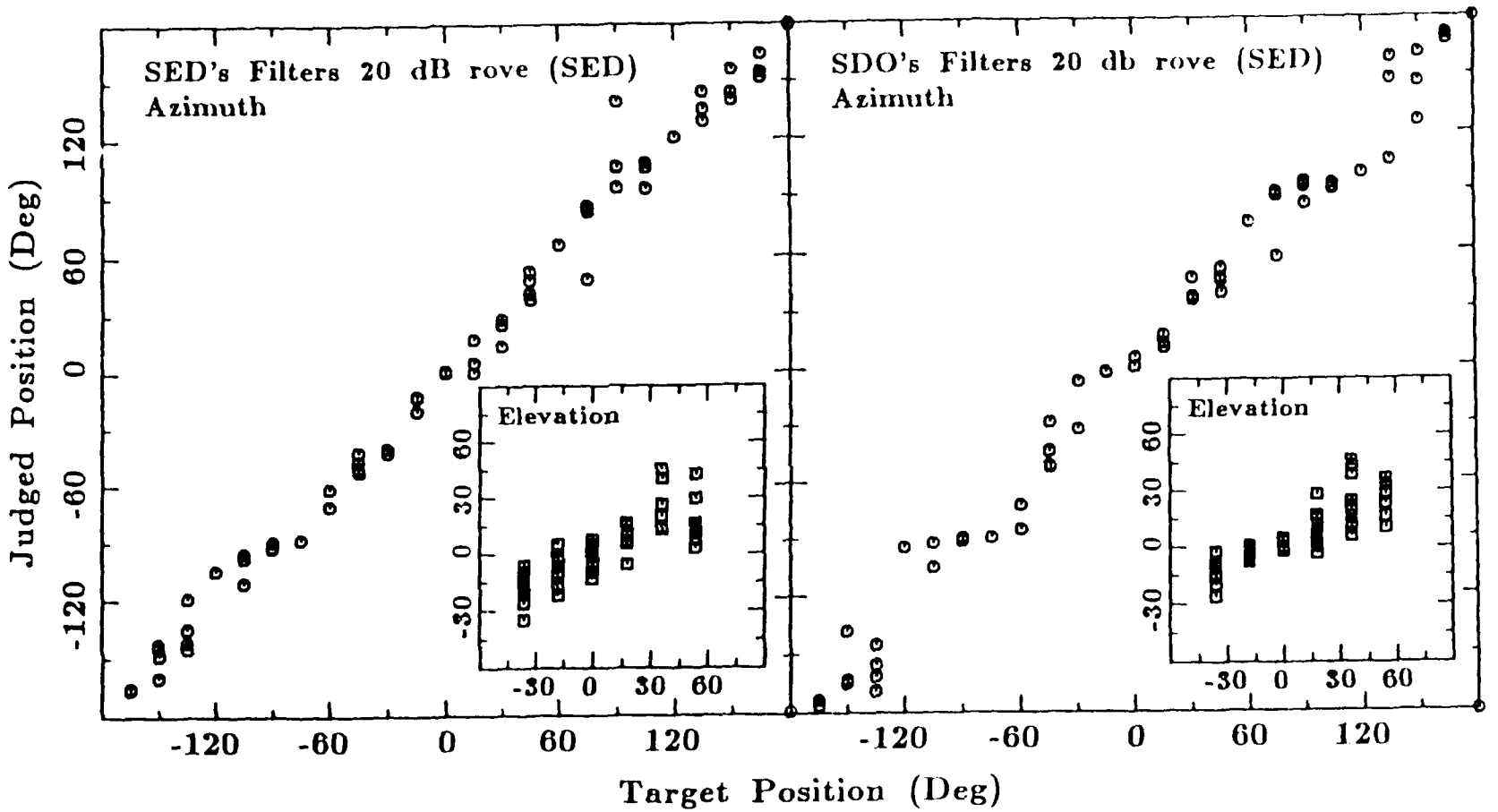


Figure 89. Same as Figure 85, except for Subject SED

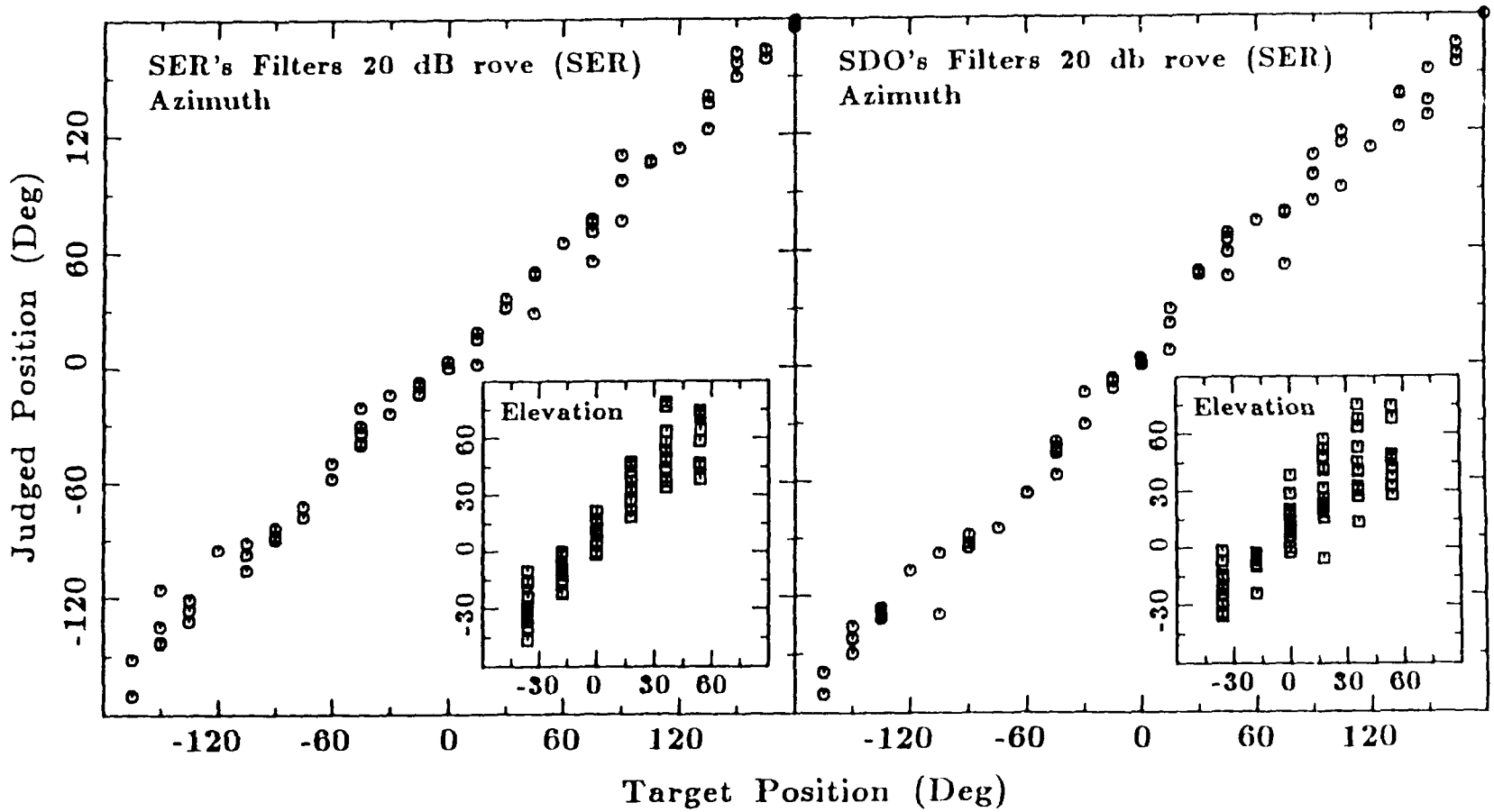


Figure 90. Same as Figure 85, except for Subject SER

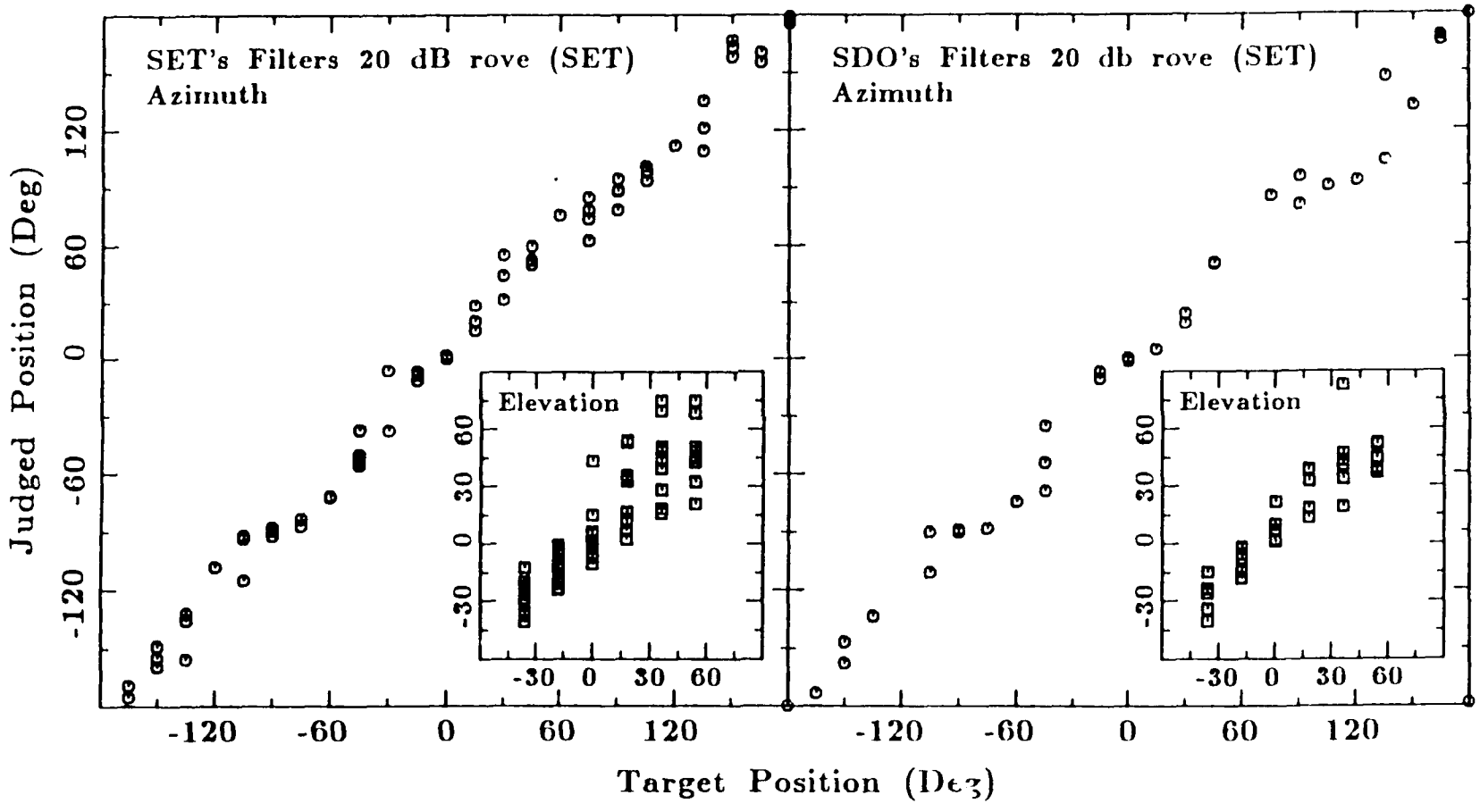


Figure 91. Same as Figure 85, except for Subject SET

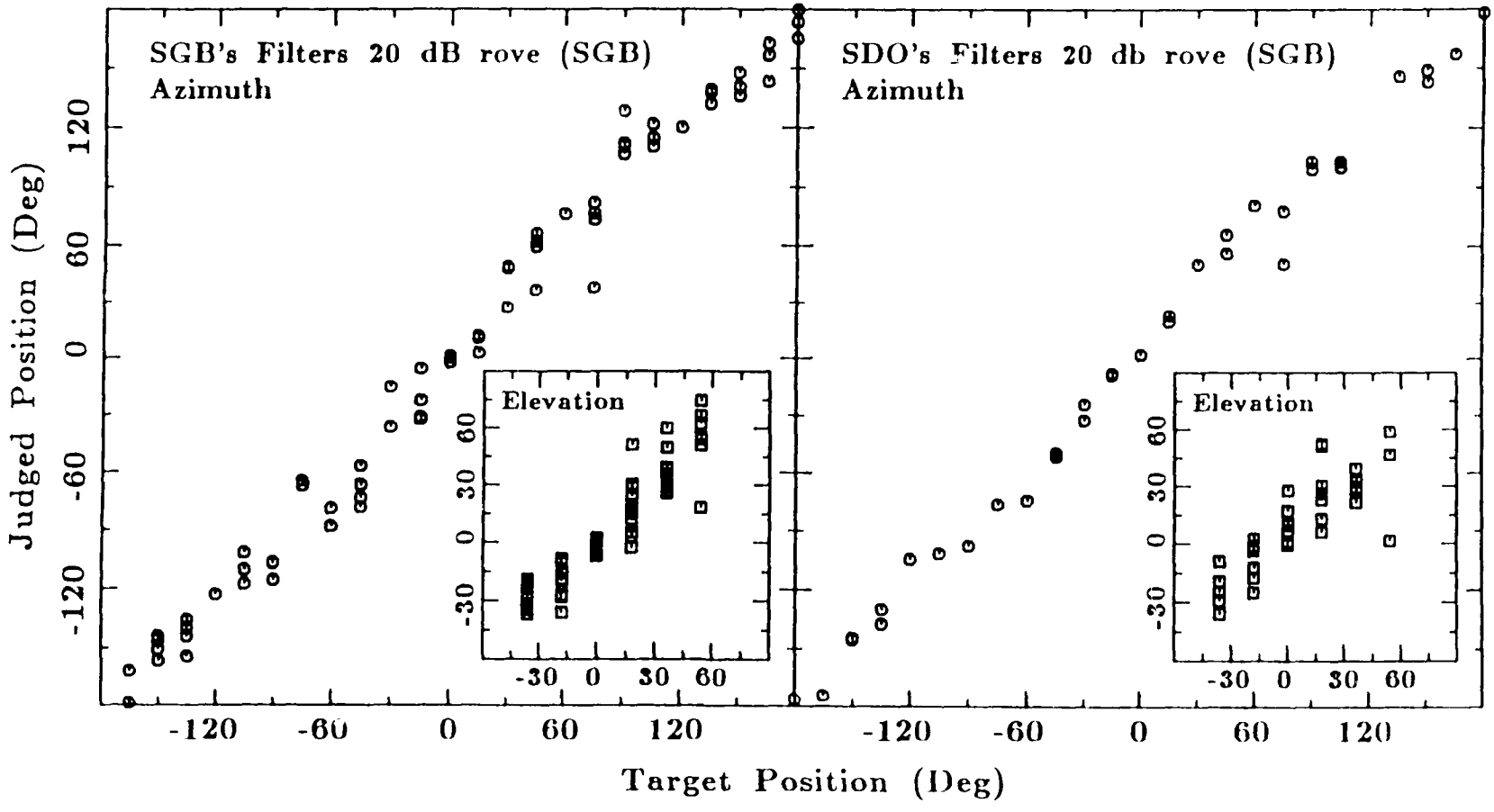


Figure 92. Same as Figure 85, except for Subject SGB

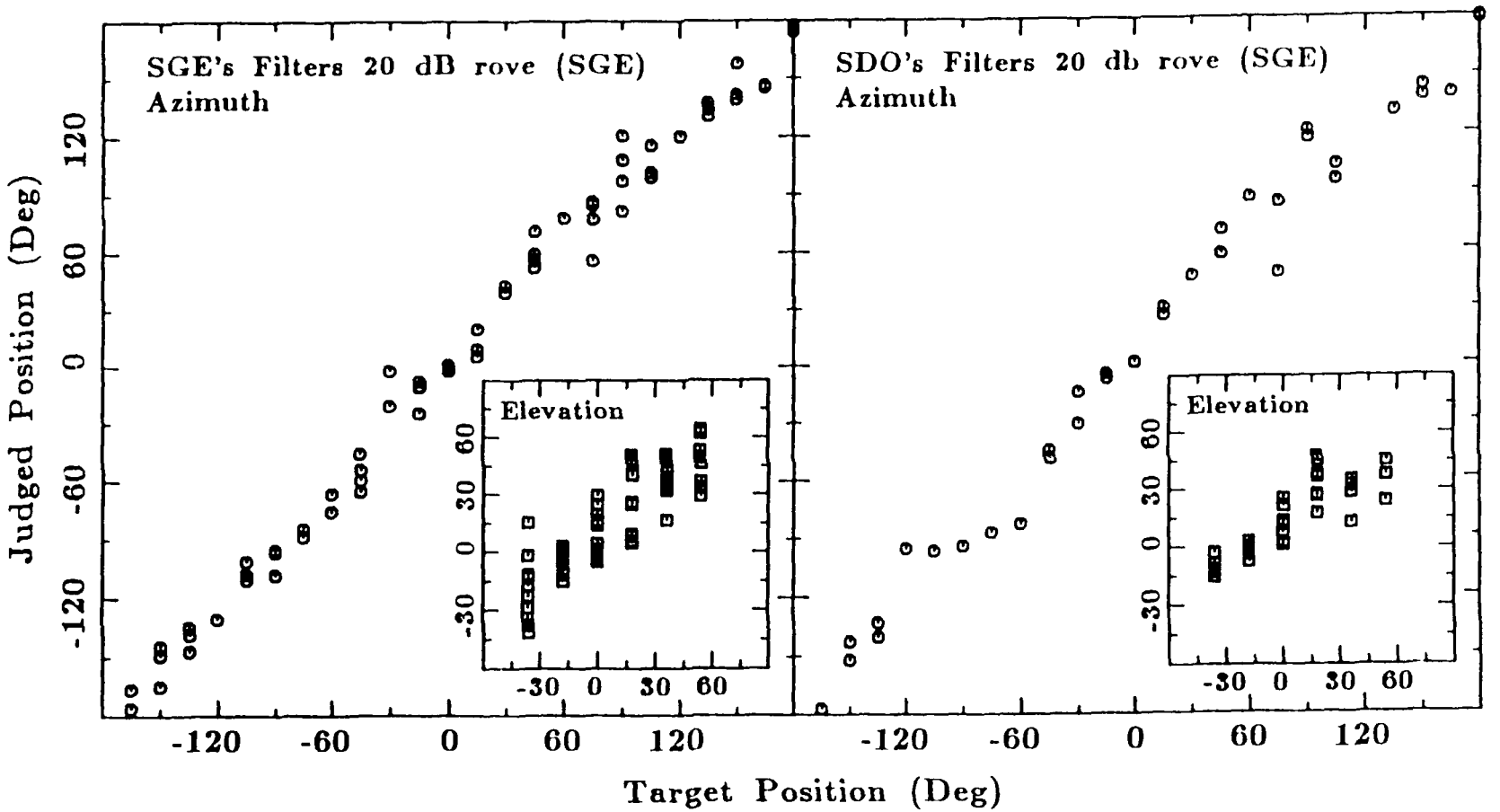


Figure 93. Same as Figure 85, except for Subject SGE

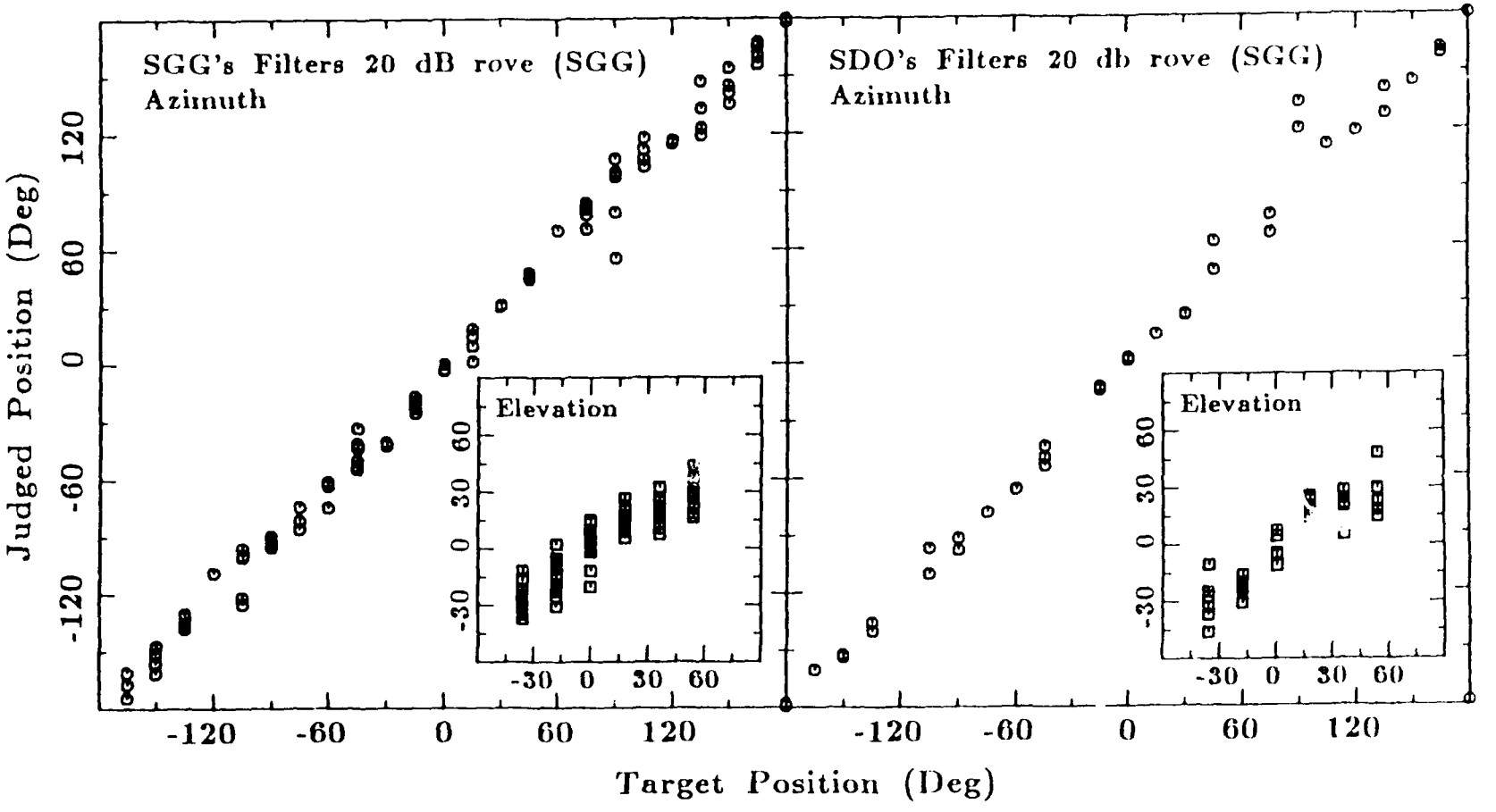


Figure 94. Same as Figure 85, except for Subject SGG

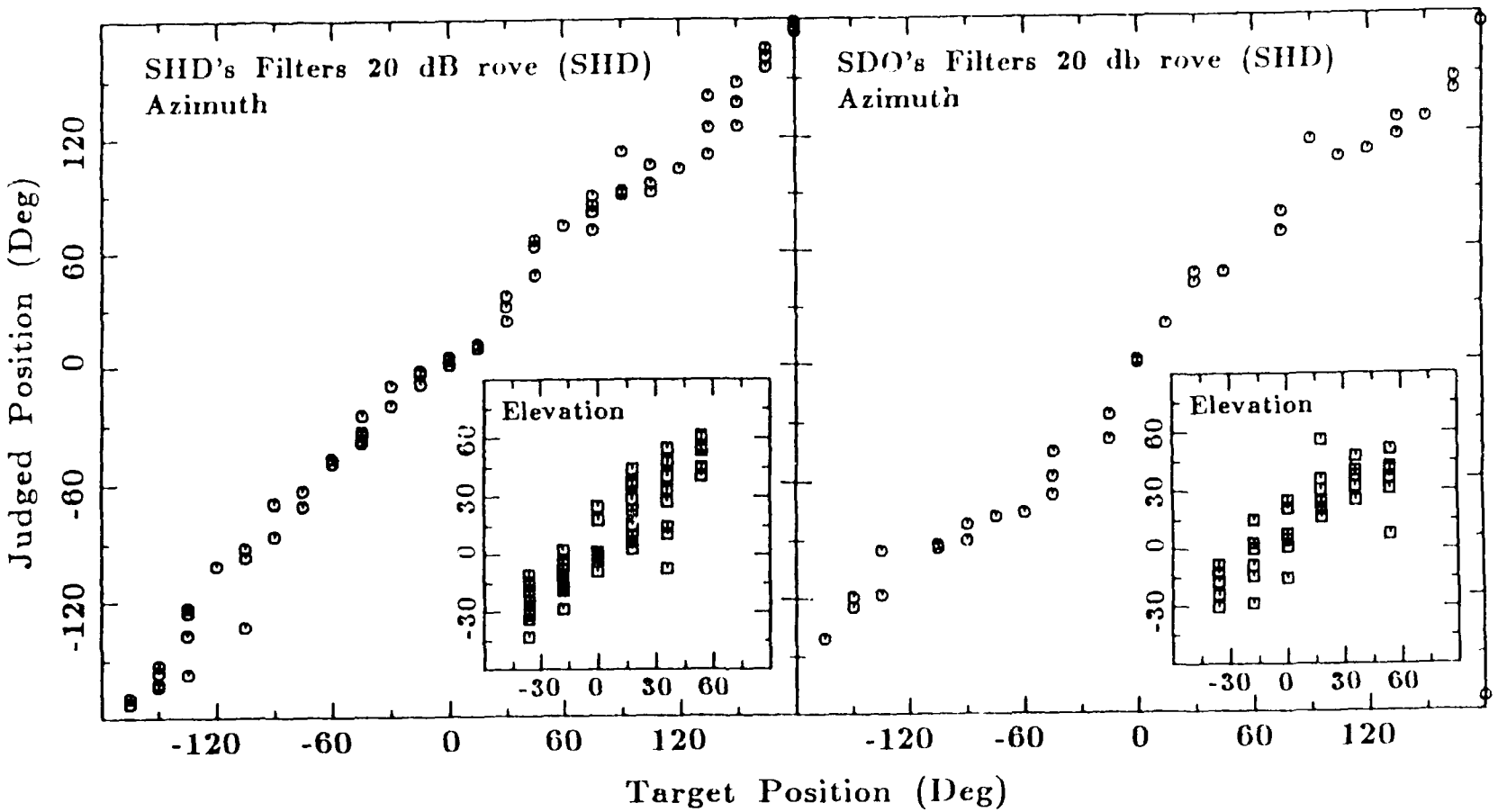


Figure 95. Same as Figure 85, except for Subject SHD

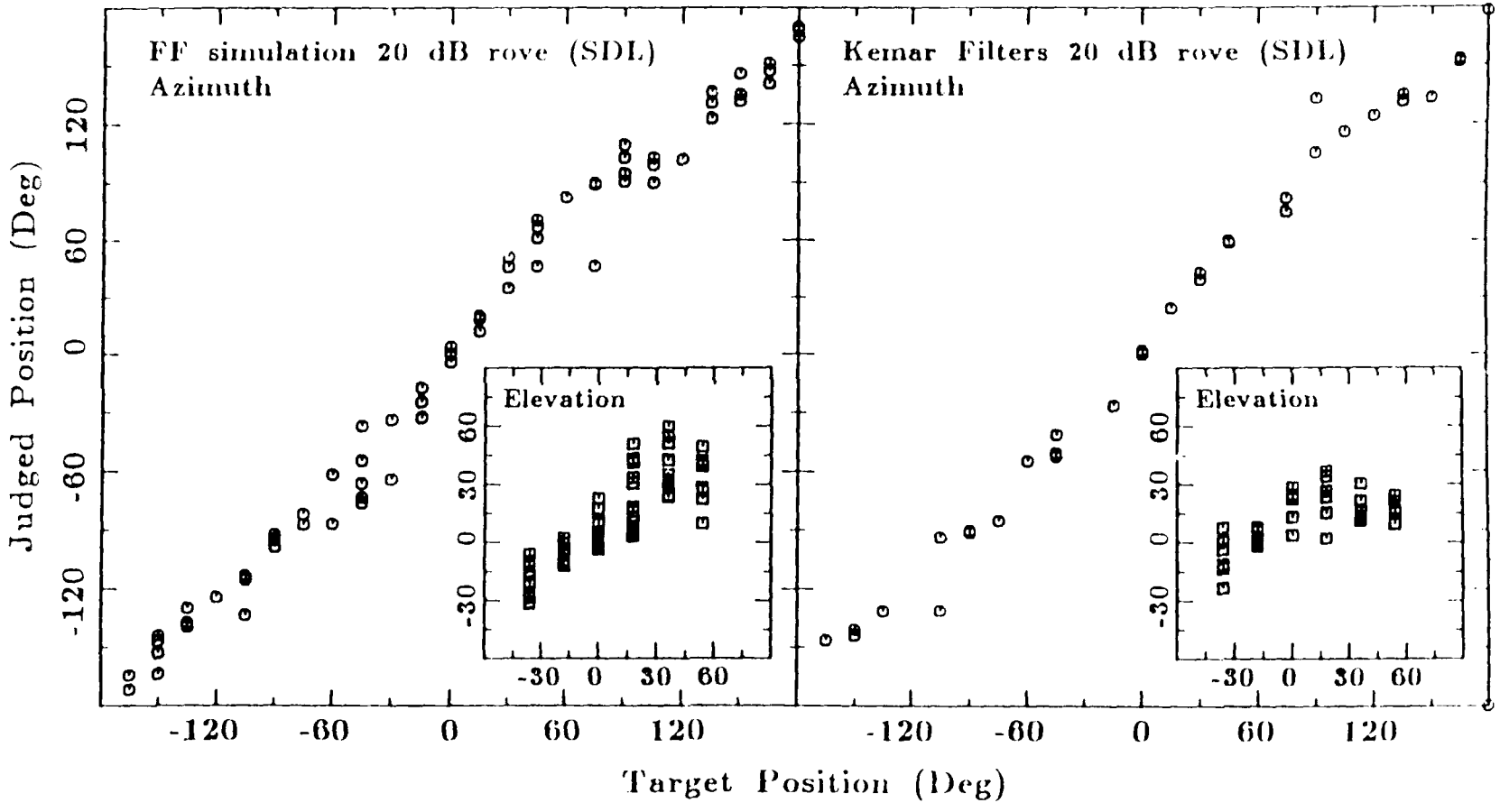


Figure 96. The right panel shows data from a condition in which SubjectSDL localized 36 stimuli synthesized from HRTFs measured from KEMAR. The format of the figure is the same as for Figure 70. The panel on the left is included for comparison and shows SDL's performance with stimuli based on SDL's own HRTFs

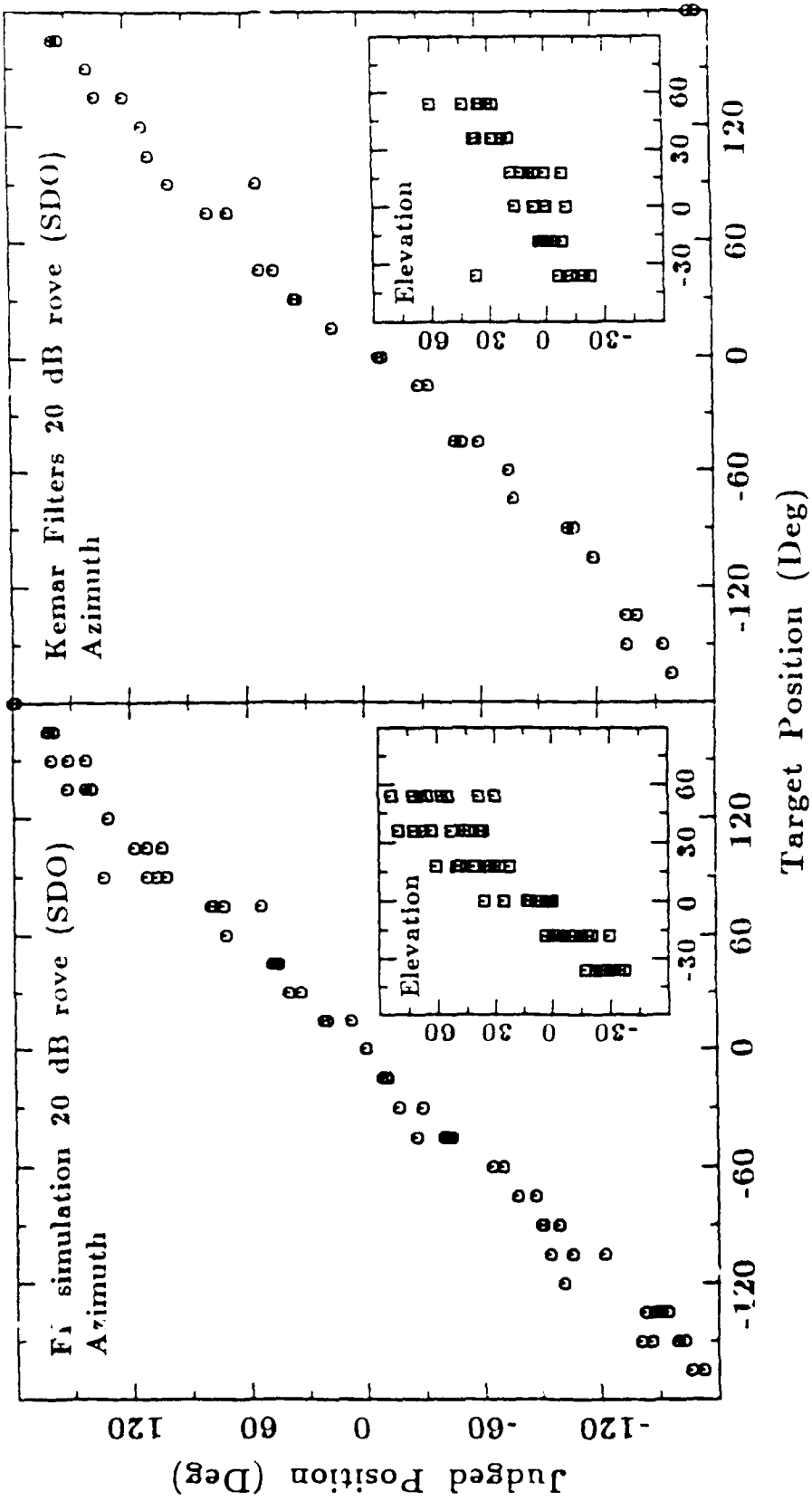


Figure 97. Same as Figure 96, except for Subject SDO

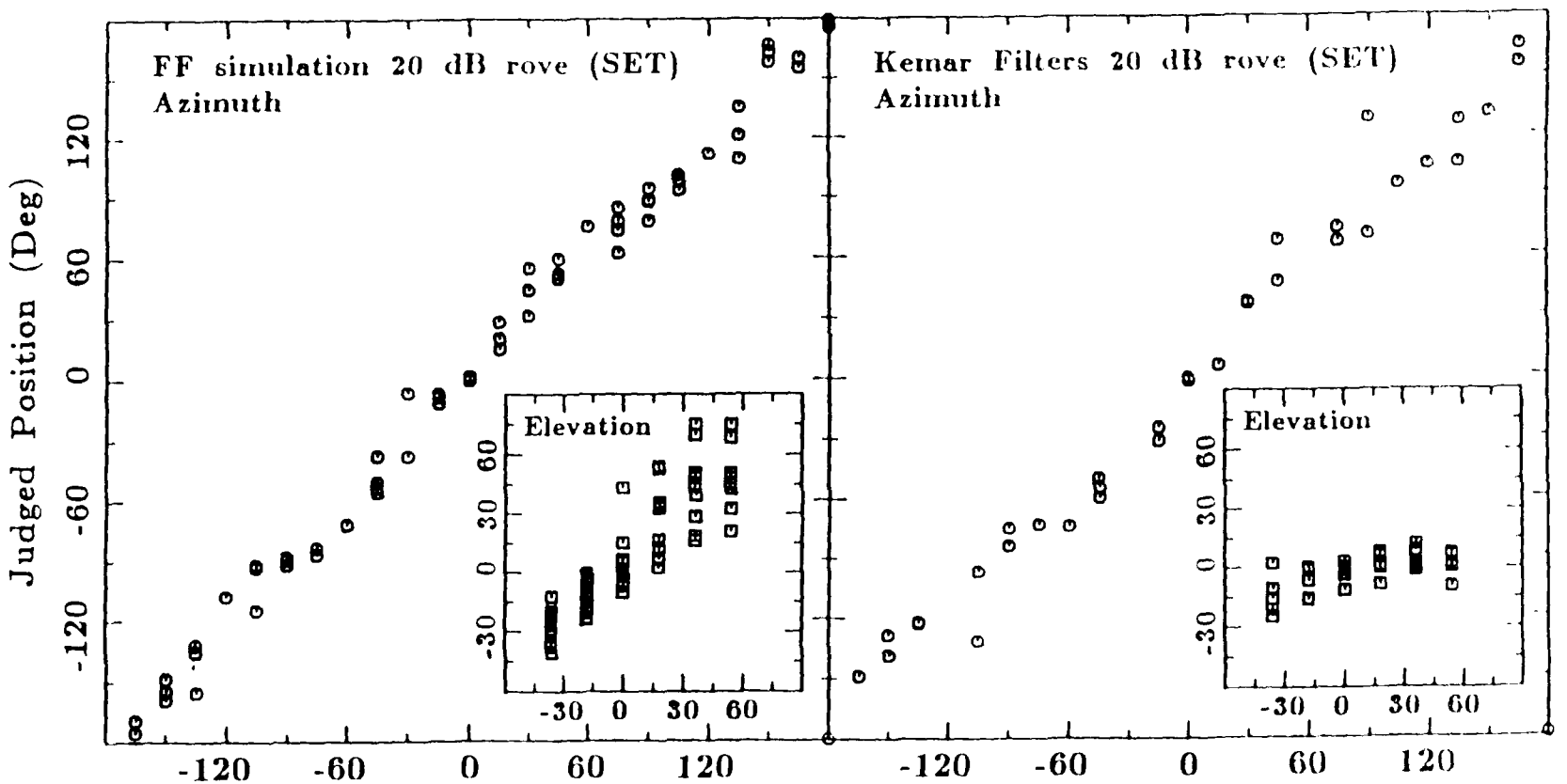


Figure 98. Same as Figure 96, except for Subject SET

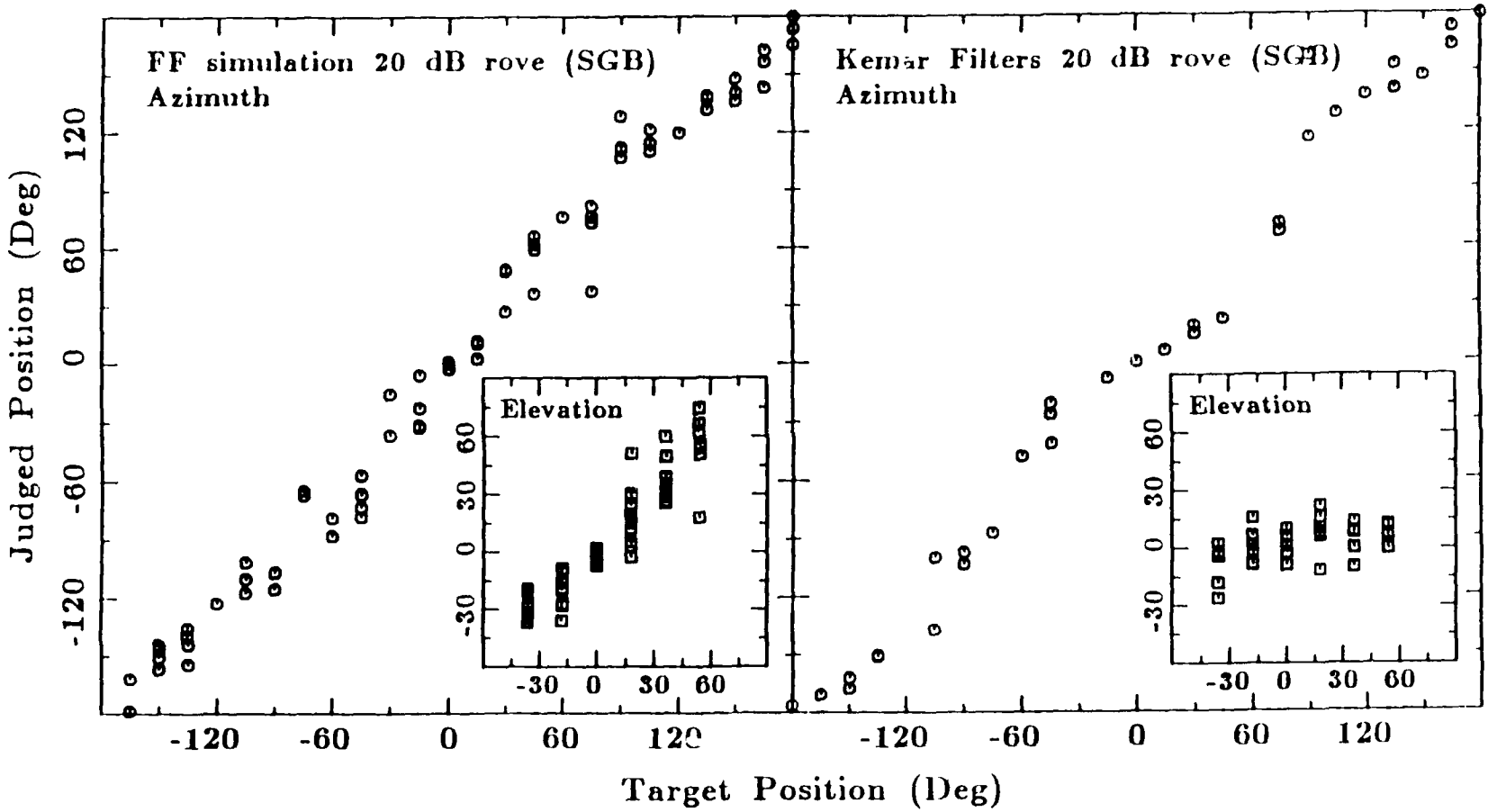


Figure 99. Same as Figure 96, except for Subject SGB

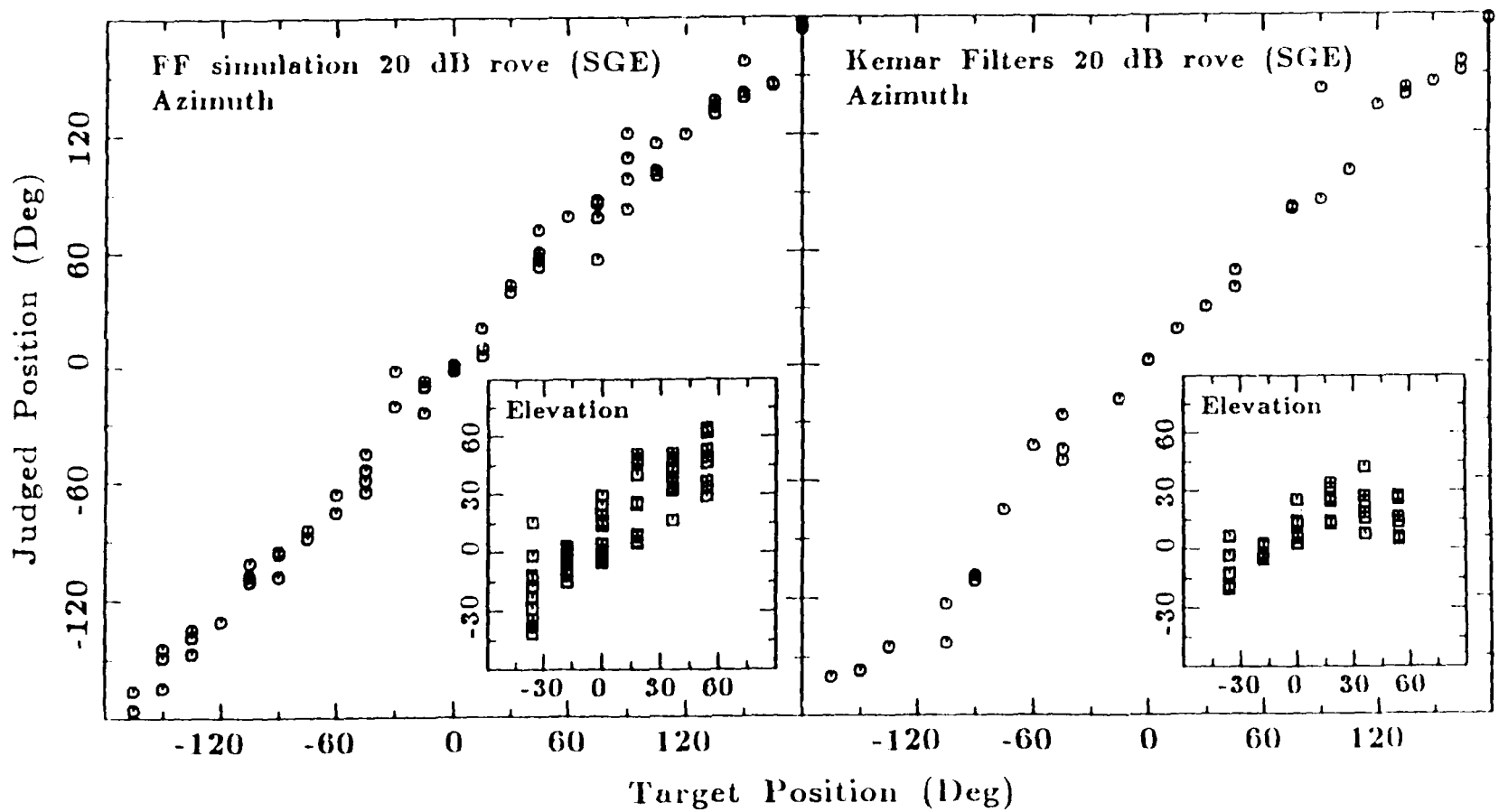


Figure 100. Same as Figure 96, except for Subject SGE

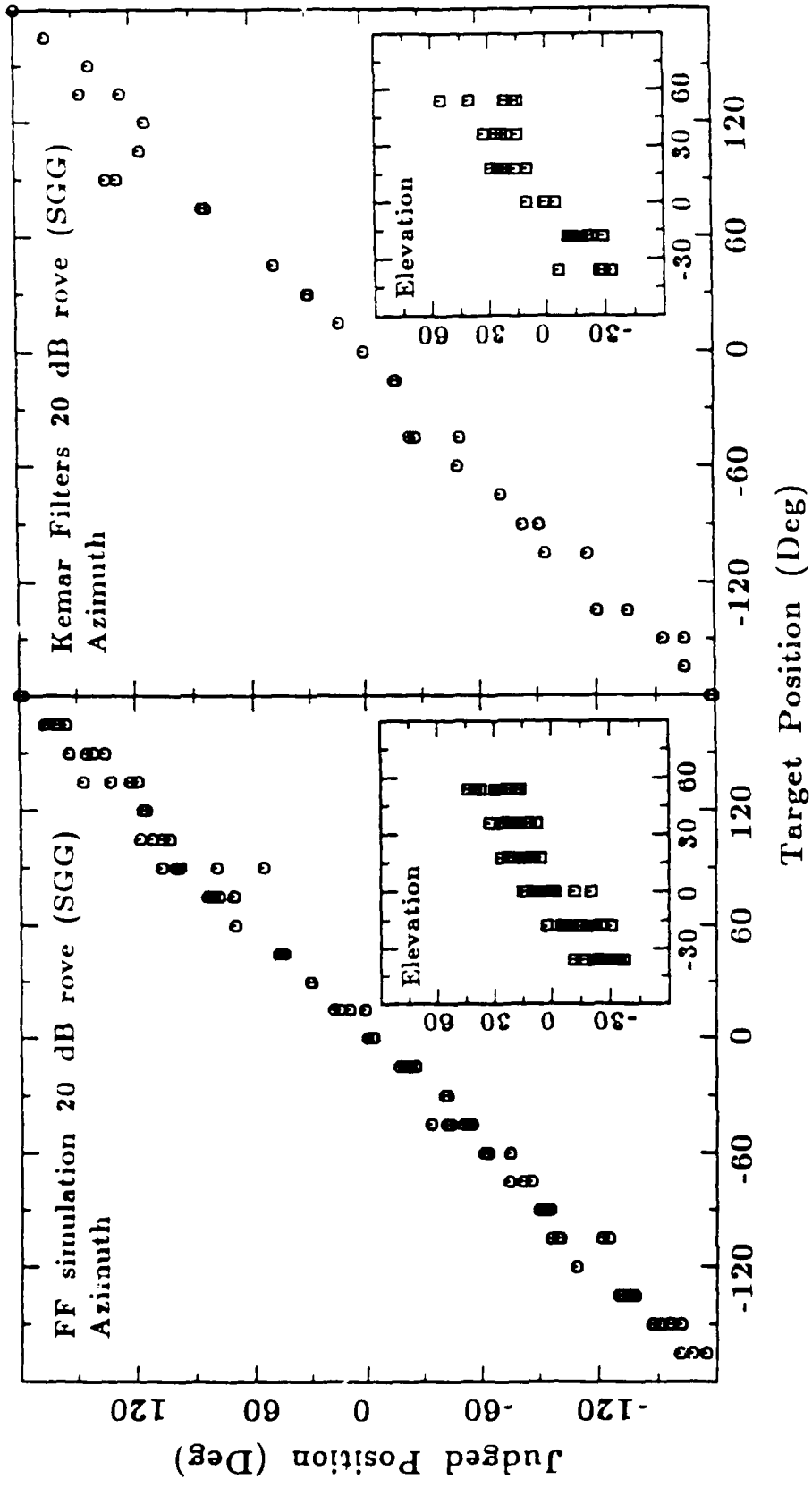


Figure 101. Same as Figure 96, except for Subject SGG

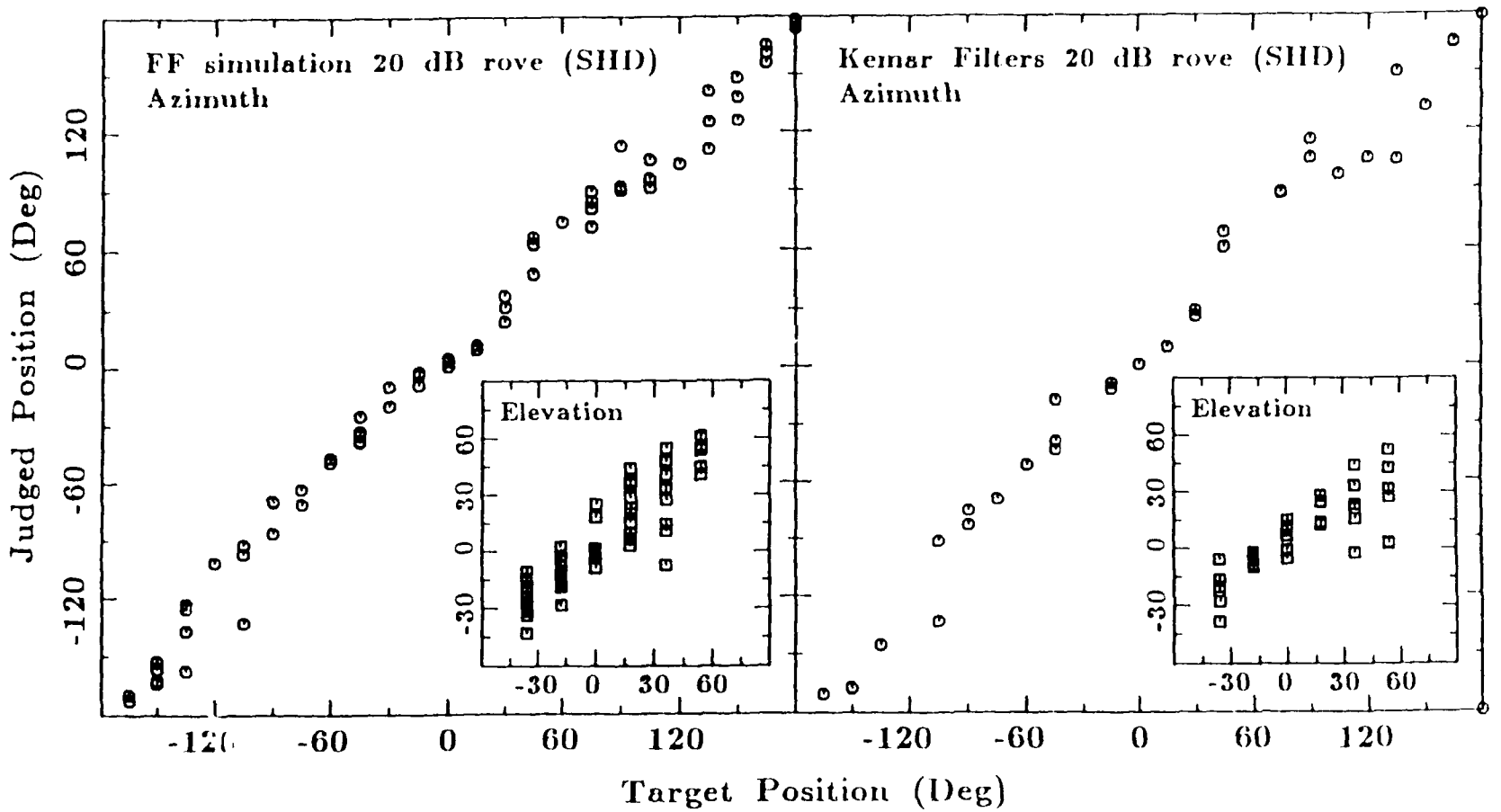


Figure 102. Same as Figure 96, except for Subject SHD

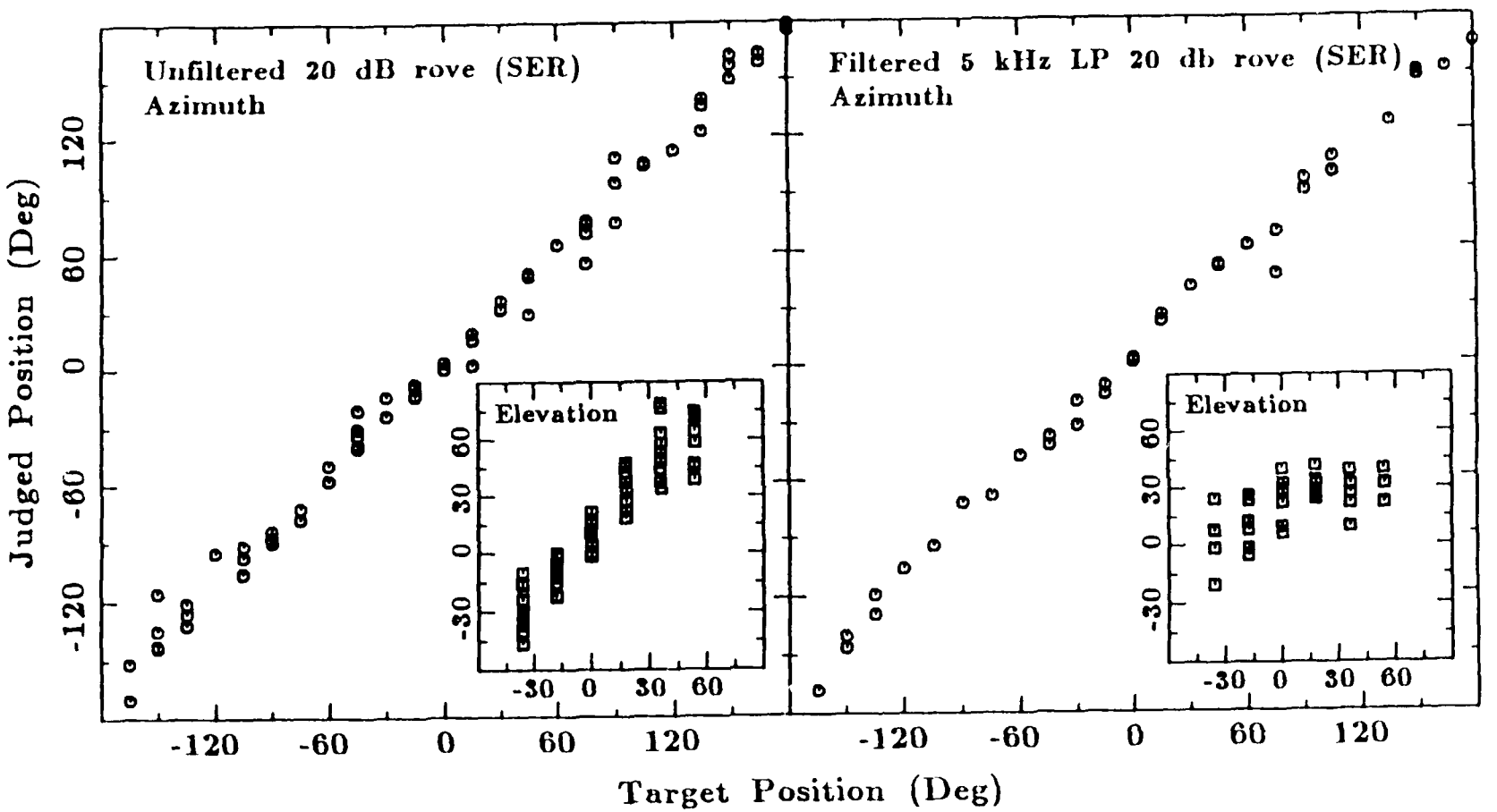


Figure 103. The right panel shows data from a condition in which Subject SER localized 36 synthesized stimuli in which the energy above 5 kHz had been removed. The format of the figure is the same as for Figure 70. The panel on the left is included for comparison and shows SER's performance with unfiltered stimuli1

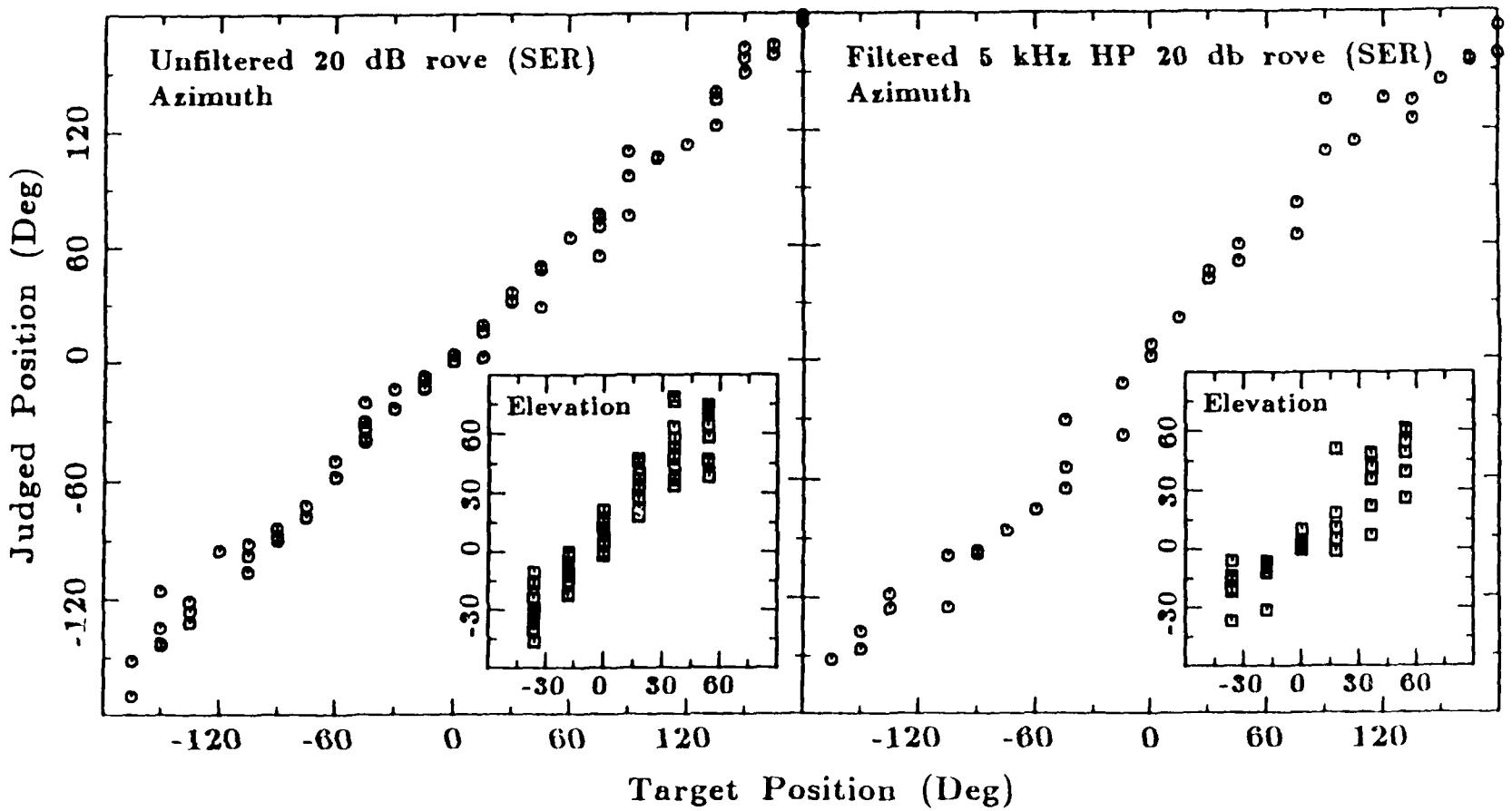


Figure 104. Same as Figure 103, except energy below 5 kHz was removed

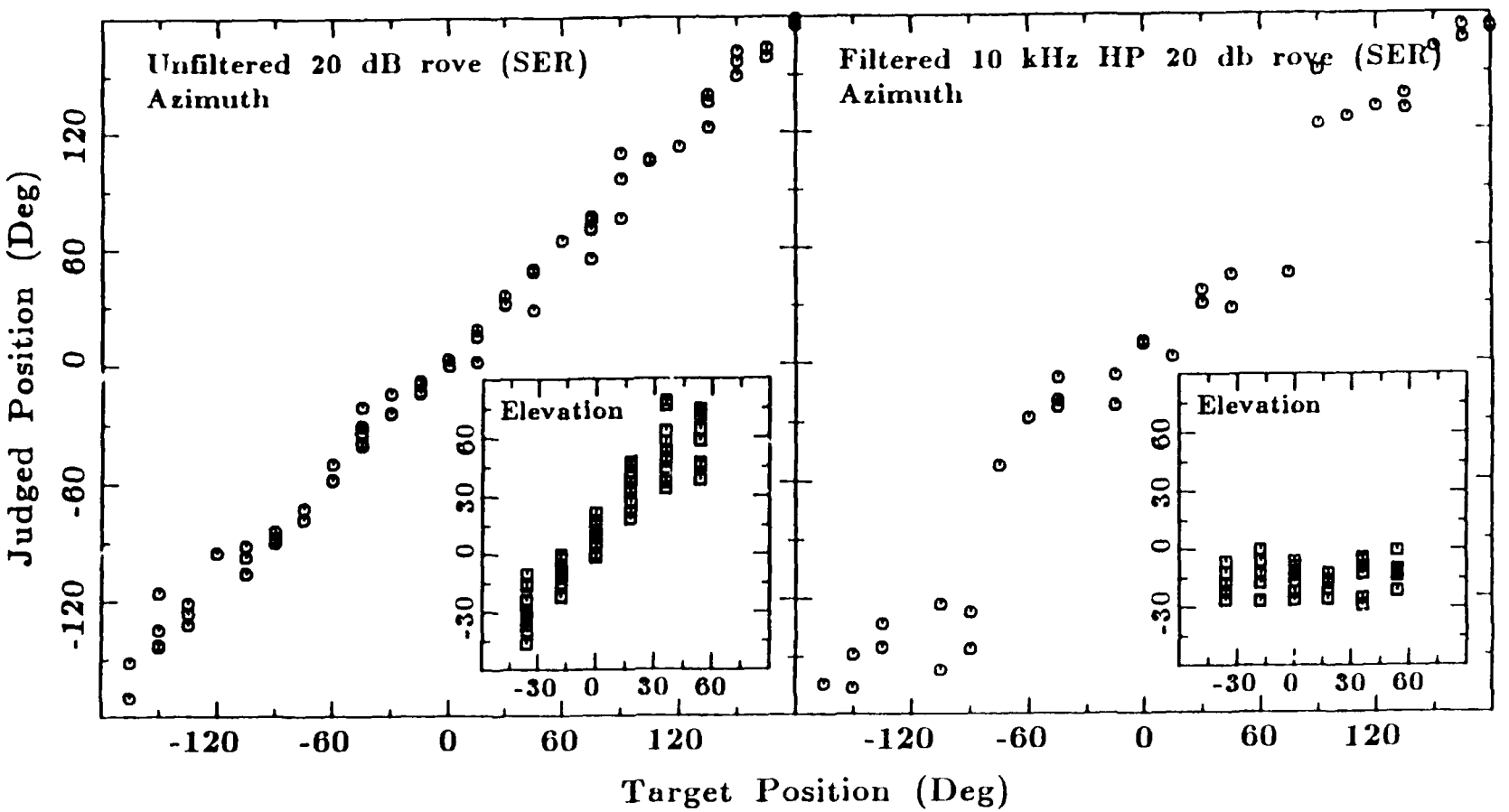


Figure 105. Same as Figure 103, except energy below 10 kHz was removed

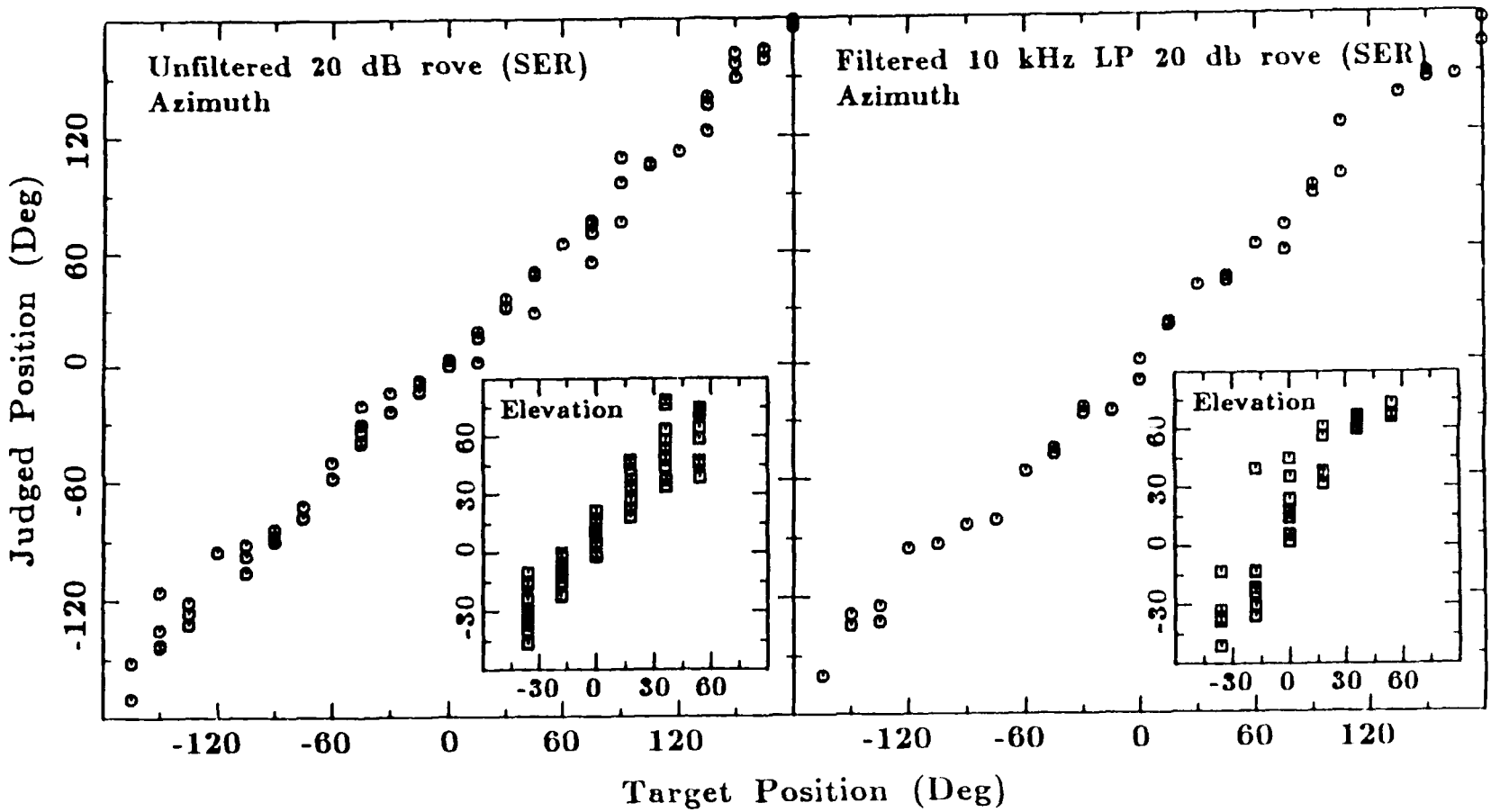


Figure 106. Same as Figure 103, except energy above 10 kHz was removed